

ÉCOLE DE TECHNOLOGIE SUPÉRIEURE
UNIVERSITÉ DU QUÉBEC

MÉMOIRE PRÉSENTÉ À
L'ÉCOLE DE TECHNOLOGIE SUPÉRIEURE

COMME EXIGENCE PARTIELLE
À L'OBTENTION DE LA
MAÎTRISE EN GÉNIE: CONCENTRATION PERSONNALISÉE
M. Sc. A.

PAR
Florent SABARLY

EVALUATION DU POTENTIEL DES REANALYSES À REPRÉSENTER LE CYCLE
HYDROLOGIQUE DE SURFACE ET DE LEUR PERFORMANCE EN MODÉLISATION
HYDROLOGIQUE DANS LA PROVINCE DE QUÉBEC AU CANADA

MONTREAL, LE 28 SEPTEMBRE 2015



Florent Sabarly, 2015



Cette licence [Creative Commons](https://creativecommons.org/licenses/by-nc-nd/4.0/) signifie qu'il est permis de diffuser, d'imprimer ou de sauvegarder sur un autre support une partie ou la totalité de cette œuvre à condition de mentionner l'auteur, que ces utilisations soient faites à des fins non commerciales et que le contenu de l'œuvre n'ait pas été modifié.

PRÉSENTATION DU JURY
CE MÉMOIRE A ÉTÉ ÉVALUÉ
PAR UN JURY COMPOSÉ DE :

Mme Annie Poulin, directrice de mémoire
Département de génie de la construction à l'École de technologie supérieure

M. Michel Baraër, président du jury
Département de génie de la construction à l'École de technologie supérieure

M. François Brissette, membre du jury
Département de génie de la construction à l'École de technologie supérieure

M. Philippe Lucas-Picher, examinateur externe
Département des sciences de la Terre et de l'atmosphère à l'Université du Québec à Montréal

IL A FAIT L'OBJET D'UNE SOUTENANCE DEVANT JURY ET PUBLIC

LE 4 SEPTEMBRE 2015

À L'ÉCOLE DE TECHNOLOGIE SUPÉRIEURE

REMERCIEMENTS

Je tiens à remercier tout d'abord Annie Poulin qui m'a donnée l'opportunité de mener à bien ce projet de recherche au sein du DRAME. Du début à la fin, tu as su m'orienter dans la bonne direction grâce à tes précieux conseils et au partage de tes connaissances.

Je remercie aussi François Brissette sans qui ma participation à la collaboration de recherche ALICE n'aurait pas été possible. Ton expertise et ton expérience dans la recherche ont su guider mes pas.

Je tiens aussi à remercier Philippe Lucas-Picher qui m'a été d'une grande aide dans la compréhension des sciences du climat, ton expertise a été déterminante dans l'aboutissement de ce projet.

Je remercie aussi Blaise Gauvin Saint-Denis d'Ouranos et Catherine Guay de l'Institut de Recherche d'Hydro-Québec pour m'avoir fourni du matériel et des informations essentiels à cette recherche.

Je pense aussi aux étudiants du DRAME pour la bonne ambiance et le soutien mutuel apporté, faire partie de cette équipe a été pour moi une expérience très enrichissante, et je vous en remercie. Je tiens aussi à te remercier particulièrement Gilles, car notre collaboration a été très motivante pour moi, merci aussi pour le partage de tes travaux et résultats.

Enfin, un grand merci à ma famille et mes amis en France et ici pour leur soutien moral malgré la distance qui nous sépare pour certains, et merci à toi Karine qui apporte de la joie et de la bonne humeur dans mon cœur depuis que je te connais!

Je souhaiterais finalement dédier ce mémoire à mon père Pierre et à mon grand-père Francis, qui ont su me transmettre (sans le vouloir sûrement!) leur passion pour les sciences de l'eau.

EVALUATION DU POTENTIEL DES REANALYSES À REPRÉSENTER LE CYCLE HYDROLOGIQUE DE SURFACE ET DE LEUR PERFORMANCE EN MODÉLISATION HYDROLOGIQUE DANS LA PROVINCE DU QUÉBEC AU CANADA

Florent SABARLY

RÉSUMÉ

La province de Québec au Canada possède une importante ressource en eau douce. Cette ressource constitue un intérêt considérable pour la province, notamment vis-à-vis du potentiel de production hydroélectrique qu'elle représente. La production d'une telle énergie nécessite de bien connaître le cycle hydrologique des régions et bassins versants d'intérêt. L'étude du cycle hydrologique s'effectue habituellement à travers l'analyse de données météorologiques et hydrologiques, mesurées grâce à des stations d'observation, et/ou à l'aide de la modélisation hydrologique. Cependant, la densité de stations d'observations est très inégale au Québec. En effet, de nombreuses stations recouvrent le sud de la province, tandis que le nord est fortement délaissé.

Toutefois, il existe des jeux de données alternatives possédant la particularité de recouvrir les régions d'intérêt au complet sous forme de grille, comme par exemple les données d'observations interpolées et les réanalyses. Cependant, les jeux de données interpolées sur grille peuvent présenter un niveau de représentativité météorologique discutable dans les régions où l'interpolation a été réalisée avec un nombre très réduit de stations. Les réanalyses sont générées de la même manière que les sorties d'un modèle de prévisions météorologiques, à la différence près que le modèle météorologique ne fait pas l'objet de mise à jour ni de modification durant le processus de génération des données. De plus, les réanalyses produisent des jeux de données historiques (et jusqu'à ce jour dans certains cas) à partir de l'assimilation de données d'observation issues de satellites, bateaux, ballons sondes, avions, bouées ou stations météo notamment.

Cette étude vise à évaluer le potentiel des réanalyses à représenter le cycle hydrologique de surface au Québec, et à évaluer leur performance en modélisation hydrologique sur des bassins versants du Québec. L'analyse du cycle hydrologique des réanalyses CFSR, NARR, MERRA et ERA-Interim est réalisée selon trois échelles spatiales différentes: (1) sur la totalité du Québec, (2) suivant les cinq régions climatiques de Bukovsky au Québec et (3) à l'échelle des bassins versants dans six régions hydrologiques de la province. Les résultats révèlent que chacune des réanalyses a ses propres forces et faiblesse. Globalement, ERA-Interim semble présenter les précipitations les plus fiables, alors que NARR montre les résultats les moins concluants. Contrairement aux trois autres réanalyses, MERRA présente un bilan hydrologique fermé. Cependant, cette réanalyse ne distribue pas, dans les proportions attendues, la précipitation en évaporation et en ruissellement. Cette dernière caractéristique est bien respectée par CFSR, toutefois, cette réanalyse présente des précipitations bien supérieures aux trois autres.

VIII

Suite à cette étude, la performance des quatre réanalyses en modélisation hydrologique a été évaluée avec le modèle global conceptuel HSAMI. La performance a été quantifiée par l'intermédiaire du Critère de Nash-Sutcliffe sur 106 bassins versants du Québec, sur les années 1979 à 2008 et selon les six régions hydrologiques québécoises. Un total de 30 calages du modèle a été réalisé, ainsi qu'une validation avec le jeu de paramètres ayant fourni le meilleur coefficient de Nash-Sutcliffe en calage sur chaque bassin versant. Les résultats indiquent globalement une bonne capacité des réanalyses à reproduire les débits observés en rivière, notamment dans les régions nordiques du Québec, avec des coefficients de Nash-Sutcliffe allant de 0.6 à 0.9. Parmi les quatre réanalyses, ERA-Interim présente les distributions de valeurs de Nash-Sutcliffe les plus élevées, bien que sa supériorité ne soit pas statistiquement significative. Par ailleurs, les précipitations peu fiables de NARR sur le Québec influencent les performances de cette réanalyse à la baisse, et révèle donc les performances les plus faibles parmi les quatre réanalyses. MERRA et CFSR font globalement jeu égal sur les six régions hydrologiques.

En conclusion, les réanalyses montrent des résultats prometteurs, tant dans leur capacité à représenter le cycle hydrologique que dans leur performance en modélisation hydrologique. Par ailleurs, en vue de valider leur applicabilité dans le cadre d'études hydrologiques, il serait pertinent d'appuyer ces résultats en comparant les données de réanalyses directement avec des données de stations d'observations, notamment dans le nord du Québec. Finalement, il serait intéressant à l'avenir de compléter l'étude de la performance en modélisation hydrologique des réanalyses par des études similaires faisant appel à d'autres types de modèles (modèles distribués et/ou à base physique) et d'autres critères d'évaluation comme le RMSE ou des critères sur les débits extrêmes.

Mots clés: Réanalyses; cycle hydrologique; modélisation hydrologique; bilan hydrologique; bassin versant

EVALUATION DU POTENTIEL DES REANALYSES À REPRÉSENTER LE CYCLE HYDROLOGIQUE DE SURFACE ET DE LEUR PERFORMANCE EN MODÉLISATION HYDROLOGIQUE DANS LA PROVINCE DU QUÉBEC AU CANADA

Florent SABARLY

ABSTRACT

The province of Quebec in Canada has abundant fresh water resources. These resources represent a considerable interest for the province, including regarding the production of hydroelectricity. The production of such energy involves understanding well the hydrological cycle in regions and river basins of interest. The study of the hydrological cycle is usually carried out through the analysis of meteorological and hydrological data, measured with weather stations, and/or with the help of hydrological modelling. However, the density of weather stations is very irregular over the province of Quebec. In fact, several stations cover the south of the province, whereas they are very few in the north.

Nevertheless, there are gridded alternative datasets that cover regions of interest entirely with gridded data, such as gridded observational datasets and reanalyses. However, observational gridded datasets may reveal questionable data in regions where interpolations have been performed with a very few weather stations. Reanalyses are generated similarly as weather forecast model outputs, with the difference that they are not updated during the production data process. Moreover, reanalyses provide datasets covering the recent past (until the present day for some of them) from the assimilation of observation data measured by different devices including satellites, boats, radio soundings, aircrafts, buoys or weather stations.

The objectives of this study are to evaluate the potential of reanalyses to represent the terrestrial hydrological cycle in the province of Quebec, and to assess their performance in hydrological modelling over river basins of the province. The analysis of the hydrological cycle of the four reanalyses CFSR, NARR, MERRA and ERA-Interim is performed according to three different spatial scales: (1) over the entire province, (2) according to the five climatic regions of Bukovsky over the province, and (3) at the watershed scale, according to six hydrological regions defined beforehand. Results show that each reanalysis has its own strengths and weaknesses. Globally, ERA-Interim seems to provide the most reliable precipitation, whereas NARR presents the less conclusive results. Contrary to the three other reanalyses, MERRA shows a closed water budget. However, this reanalysis does not distribute the precipitation into runoff and evaporation according to the expected amounts of water. This last aspect is well respected by CFSR, but this reanalysis also reveals the highest precipitation over the province of Quebec, among the four reanalyses.

Following this study, the performance of the four reanalyses in hydrological modelling has been evaluated with the global conceptual HSAMI model. The performance has been assessed using the Nash-Sutcliffe criteria over 106 river basins of the province of Quebec,

during the 1979-2008 period, and according to the six hydrological regions. Thirty (30) calibrations have been performed, as well as one validation with the parameters that provided the best Nash-Sutcliffe criteria in calibration for each river basin. Results indicate globally a good capacity of reanalyses to reproduce the observed river discharges, including in northern regions of the province, where the Nash-Sutcliffe criteria varies from 0.6 to 0.9. Among the four reanalyses, ERA-Interim presents the highest Nash-Sutcliffe value distributions, although its superiority is not statistically significant. On the other hand, the questionable precipitation of NARR over the province of Quebec influences the performance of this reanalysis downward, and thus reveals the lowest performance among the four reanalyses. MERRA and CFSR show quite similar results over the six hydrological regions of the province.

In conclusion, reanalyses show promising results both in their potential to represent the hydrological cycle and in their performance in hydrological modelling. Moreover, in order to validate their applicability regarding hydrological studies, it is recommended to strengthen these results with additional comparisons between reanalyses data and weather stations data, especially over the northern regions of the province. Finally, it would be interesting in future studies to assess the hydrological performance of reanalyses using different types of hydrological models (distributed models, and/or physical basis) and other assessment criteria, as the RMSE or criteria based on extreme events.

Keywords: Reanalyses; hydrological cycle; hydrological modelling; water budget; river basin

TABLE DES MATIÈRES

	Page
INTRODUCTION	1
CHAPITRE 1 REVUE DE LITTÉRATURE	5
1.1 Les réanalyses	5
1.1.1 Qu'est-ce qu'une réanalyse?	5
1.1.2 Naissance et évolution des réanalyses	6
1.1.3 Études comparatives des réanalyses	8
1.2 La modélisation hydrologique	11
1.2.1 Qu'est-ce qu'un modèle hydrologique?	11
1.2.2 Modélisation hydrologique et réanalyses.....	13
CHAPITRE 2 USE OF REANALYSIS DATA TO ASSESS THE TERRESTRIAL BRANCH OF THE WATER CYCLE: A CASE STUDY OF FOUR DATASETS OVER QUEBEC, CANADA	17
2.1 Abstract	17
2.2 Introduction.....	18
2.3 Data and methods.....	21
2.3.1 Reanalysis datasets.....	22
2.3.2 Observational datasets	24
2.3.3 Methodology	26
2.3.3.1 Long-term mean of the hydrological cycle components over the province of Quebec	27
2.3.3.2 Hydrological cycle components over the climatic regions	28
2.3.3.3 Hydrological cycle components within river basins	29
2.4 Results.....	31
2.4.1 Long-term mean of the hydrological cycle components over the province of Quebec	31
2.4.2 Hydrological cycle components over the climatic regions	35
2.4.3 Hydrological cycle components within river basins	39
2.5 Concluding remarks	45
2.6 Acknowledgements.....	48
CHAPITRE 3 ÉTUDE DE LA PERFORMANCE DES RÉANALYSES EN MODÉLISATION HYDROLOGIQUE	49
3.1 Méthodologie	49
3.1.1 Préparation des données.....	49
3.1.1.1 Bassins versants de $(cQ)^2$	50
3.1.1.2 Réanalyses et données d'observations sur grille	55
3.1.2 Calage et validation du modèle HSAMI	56
3.2 Résultats.....	58
3.2.1 Calage	58

3.2.2	Validation.....	61
3.2.3	Analyses complémentaires.....	66
3.3	Discussions et conclusions.....	69
SYNTHÈSE ET CONCLUSION		73
ANNEXE I	LISTE DES FIGURES RELATIVES AU TEST STATISTIQUE MULTIPLE DE COMPARAISON DE KRUSKAL-WALLIS EFFECTUÉ ENTRE LES DISTRIBUTIONS DES NSE DE CHAQUE JEU DE DONNÉES SUR LES SIX RÉGIONS HYDROLOGIQUES DU QUÉBEC.....	77
ANNEXE II	FIGURES DES MOYENNES SUR LA PÉRIODE 1979 - 2008 DES TEMPÉRATURES MINIMALES ET MAXIMALES DE NRCAN, CFSR, ERA-INTERIM ET MERRA SUR LES TERRITOIRES DE L'EST DU CANADA	83
LISTE DE RÉFÉRENCES BIBLIOGRAPHIQUES.....		85

LISTE DES TABLEAUX

	Page
Tableau 2.1	Main properties summary of reanalysis datasets22
Tableau 2.2	Correlation between daily time series of spatially averaged precipitation of CFSR, ERA-Interim, MERRA and NARR versus NRCAN over the 11 selected river basins42
Tableau 3.1	Performances (NSE) du modèle HSAMI avec NRCAN, NARR, ERA-Interim, CFSR et MERRA en validation sur 12 bassins versants du Québec. Les moyennes par bassin versant et par jeu de données sont indiquées respectivement dans la dernière colonne et dernière ligne.....68

LISTE DES FIGURES

	Page
Figure 0.1	Cycle de l'eau.....1
Figure 2.1	Location of the 956 Environment Canada weather stations used to generate the daily NRCAN dataset (Hutchinson et al., 2009). The stations are represented only over the eastern territories of Canada.....25
Figure 2.2	(a) Climatic regions of Bukovsky for CFSR, and (b) Hydrologic regions and selected river basins (with red contours) from (cQ) ²31
Figure 2.3	Long-term mean precipitation, evaporation and runoff (mm day ⁻¹) of CFSR, ERA-Interim, MERRA and NARR over the province of Quebec33
Figure 2.4	Long-term mean relative water balances using <i>B</i> values (% precip) from CFSR, ERA-Interim, MERRA and NARR over the province of Quebec35
Figure 2.5	Mean annual cycle of precipitation, evaporation and runoff (in mm day ⁻¹), and annual mean relative water balance (in % of precip), from 1979 to 2008, of MERRA, ERA-Interim, NARR and CFSR, with the addition of NRCAN for the case of precipitation for the five climatic regions from Bukovsky (2011)38
Figure 2.6	Mean annual precipitation (mm day ⁻¹) from CFSR, ERA-Interim, MERRA, NARR and NRCAN, from 1979 to 2008 over the 11 selected river basins. For each river basin, the total area is indicated in square kilometers, and the hydrologic region is indicated on the bottom left corner40
Figure 2.7	Mean precipitation of NARR (mm day ⁻¹) for the periods 1979-2002 (left) and 2003-2008 (right), over the province of Quebec.....41
Figure 2.8	Mean annual runoff (mm day ⁻¹) from CFSR, ERA-Interim, MERRA, NARR and NRCAN, from 1979 to 2008 over the 11 selected river basins. Annual time series are limited to the available streamflow observations43
Figure 2.9	Annual E/P ratio (evaporation divided by precipitation) from CFSR, ERA-Interim, MERRA, NARR and observational estimates from 1979 to 2008 over the 11 selected river basins. Annual time series are

	limited to the available streamflow observations that were used to estimate E from observations.....	45
Figure 3.1	Localisation des 306 bassins versants de $(cQ)^2$ et leurs gestionnaires.....	50
Figure 3.2	Schéma de l'algorithme de filtrage des bassins versants de $(cQ)^2$. Les critères des filtres indiqués dans les boîtes concernent les bassins versants qui sont conservés.....	52
Figure 3.3	Localisation des 106 bassins versants conservés après filtrage, selon les régions hydrologiques du Québec. Les 12 bassins versants dont les cycles annuels moyens des débits simulés et observés sont analysés dans la section 3.2.3, sont indiqués par un contour rouge	55
Figure 3.4	Distribution par région hydrologique des coefficients de Nash-Sutcliffe (en ordonnée) pour les jeux de données MERRA, CFSR, ERA-Interim, NARR et NRCAN. Il y a deux boîtes à moustaches pour chaque jeu de données: la première correspond à la distribution des NSE pour les 30 calages de chaque bassin, la seconde correspond à la distribution des meilleurs NSE pour chaque bassin (une valeur par bassin). Le titre de chaque sous-figure indique le nombre de bassins versants dans chaque région hydrologique, ainsi que la superficie moyenne des bassins versants.....	59
Figure 3.5	Séries des débits observés journaliers sur les bassins versants (a) Lac St-François, (b) Gouin, (c) Mountain Chute et (d) Madawaska-Arnrior (en $m^3 s^{-1}$). La région hydrologique est indiquée en haut à gauche et la superficie de chaque bassin versant dans le titre des sous-figures.....	61
Figure 3.6	Distribution par région hydrologique des coefficients de Nash-Sutcliffe (en ordonnée) en validation pour les jeux de données MERRA, CFSR, ERA-Interim, NARR et NRCAN. Il y a deux boîtes à moustaches pour chaque jeu de données: la première correspond aux NSE ayant été produits par les jeux de paramètres générant le meilleur NSE parmi les 30 calages. La seconde correspond aux meilleurs NSE parmi les 30 validations.....	64
Figure 3.7	Hydrogrammes moyens annuels des débits simulés par NRCAN, NARR, ERA-Interim, CFSR et MERRA et des débits observés sur 12 bassins versants du Québec. Les débits sont indiqués en ordonnées ($m^3 s^{-1}$) et les mois en abscisse	67
Figure 3.8	Distribution en nuage de points des coefficients de Nash-Sutcliffe (en ordonnée) des 99 bassins versants sélectionnés pour les jeux de données NRCAN, NARR, ERA-Interim, CFSR et MERRA. L'abscisse représente les superficies classées dans l'ordre croissant des bassins	

versants, allant de 10 à 69191 km ² (de gauche à droite sur chaque sous-figure)	69
--------------------------------------------------------------------------------------------------	----

LISTE DES ABRÉVIATIONS, SIGLES ET ACRONYMES

CEHQ	Centre d'Expertise Hydrique du Québec
CEPMMT	Centre Européen de Prévision Météorologique à Moyen Terme
CFSR	Climate Forecast System Reanalysis
CMAES	Covariance Matrix Adaptation Evolution Strategy
(cQ) ²	Impact des Changements Climatiques sur l'Hydrologie (Q) au Québec
DOE	Department Of Energy
EB	East Boreal
ECMWF	European Center for Medium-Range Forecasts
ERA-15	ECMWF Re-Analysis 15 years
ERA-40	ECMWF Re-Analysis 40 years
ERA-Interim	ECMWF Re-Analysis Interim
ETA	East Tundra
ETU	East Taïga
HQ	Hydro-Québec
IAU	Incremental Analysis Update
JMA	Japan Meteorological Agency
JRA-25	Japanese Re-Analysis 25 years
GL	Great Lakes
MERRA	Modern Era retrospective analysis for Research and Applications
MOPEX	Model Parameter Estimation Experiment
NA	North Atlantic

NARCCAP	North American Regional Climate Change Assessment Program
NARR	North American Regional Reanalysis
NASA	National Aeronautics and Space Administration
NCAR	National Center for Atmospheric Research
NCEP	National Centers for Environmental Prediction
NRCAN	natural Resources Canada
NSE	Coefficient de Nash-Sutcliffe (Nash-Sutcliffe Efficiency)
RTA	Rio Tinto Alcan
3D-VAR	Three Dimensional Variational scheme
4D-VAR	Four Dimensional Variational scheme

INTRODUCTION

Le cycle hydrologique est le cycle complet des phénomènes subis par l'eau depuis la vapeur d'eau atmosphérique, en passant par l'état liquide ou solide sous forme de précipitations, pour s'écouler sur ou sous la terre jusqu'à la mer, et finalement retourner à la forme de vapeur d'eau atmosphérique par l'action de l'évaporation et de la transpiration (Association québécoise des techniques de l'eau, 1981). Le cycle hydrologique (Figure 0.1) est donc à la base de la régulation des masses d'eau, processus essentiel au maintien et à l'évolution du vivant. De la même manière, le développement des activités anthropiques est fortement lié à l'évolution de l'eau dans le milieu naturel.



Figure 0.1 Cycle de l'eau
Tirée de John M. Evans (2013)

En effet, connaître les débits de rivières ou la précipitation permet par exemple de dimensionner des ouvrages d'art tels que des ponts, des digues ou des routes. De même, pouvoir estimer les réserves souterraines ou de surface en eau douce permet de gérer plus

efficacement les ressources hydriques à des fins de production d'hydroélectricité, ainsi que les réserves d'eau destinée à l'approvisionnement en eau potable ou à l'agriculture.

La province de Québec au Canada s'étale sur une surface de plus de 1,5 millions de kilomètres carrés et possède l'une des plus importantes réserves en eau douce du monde. En effet, ce territoire, dont une bonne partie demeure relativement inoccupée, est parsemé de milliers de lacs et rivières, essentiels au maintien de la biodiversité locale. Par ailleurs, cette importante réserve d'eau représente un potentiel considérable pour la production d'électricité d'origine renouvelable.

L'évaluation du potentiel hydroélectrique d'un système hydrique nécessite une bonne connaissance du cycle hydrologique local et de son évolution. Pour cela, il est nécessaire d'analyser les données météorologiques, telles que la température et la précipitation, ainsi que les données hydrologiques, telles que les débits en rivière, l'humidité du sol, etc. Cependant, les données d'observation habituellement colligées par des stations automatisées ou manuelles se font rares dans les régions reculées du Québec, notamment dans le nord, où la densité spatio-temporelle de relevés météorologiques est très faible, voire nulle.

Néanmoins, des moyens alternatifs peuvent être envisagés de façon à pallier ce manque d'information. En effet, il existe des jeux de données météorologiques générés synthétiquement sur une grille spatiale régulière à partir de l'interpolation des données aux stations. Ces jeux de données permettent de couvrir la totalité d'un territoire et sur une période temporelle donnée en fournissant des données de température et de précipitation uniquement, et sont susceptibles de révéler des écarts par rapport à la réalité, lorsque les points de grille sont interpolés à partir d'un nombre très faible de stations d'observations. Une autre possibilité serait d'utiliser des données « historiques » générées par un modèle de prévision météorologique. Un tel modèle génère des prévisions en trois dimensions sous forme de données sur grille régulière (par exemple précipitation, température, évaporation, humidité de l'air, etc), par l'intermédiaire d'un processus d'assimilation de données d'observation, ce qui force le modèle à produire un état de l'atmosphère le plus proche

possible de la réalité. Les observations assimilées regroupent différents types de données comme la direction et l'intensité du vent, la présence de neige, les flux radiatifs, la température, etc., et peuvent aussi inclure des observations au sol. Ainsi une réanalyse pourrait révéler des résultats intéressants, notamment vis-à-vis des données d'observations interpolées sur grille, dans les régions où les stations d'observations sont absentes.

Les données météorologiques du passé récent sont utiles pour la recherche en science du climat, notamment pour l'évaluation des changements climatiques. Or, depuis l'apparition des premiers modèles de prévision météorologiques, ces derniers ont fait l'objet de nombreuses mises à jour et améliorations, ce qui a entraîné l'apparition d'écarts significatifs dans la qualité et la résolution spatio-temporelle entre les données générées les plus anciennes et les plus récentes. Depuis les années 90, différents organismes de recherche en climat et en météorologie ont donc entrepris la création de jeux de données qui remédient à ce problème. Ces jeux de données, appelés réanalyses, sont générés de la même manière que pour la prévision météorologique, à ceci près que le modèle produisant la réanalyse et la méthode d'assimilation des données observées restent inchangés durant la génération de la réanalyse. D'autre part, les réanalyses fournissent habituellement un large éventail de variables atmosphériques, notamment les composantes du cycle hydrologique de surface: précipitation, évaporation et ruissellement. Par conséquent, les réanalyses pourraient se montrer très utiles dans le cadre de l'étude du cycle hydrologique de surface, notamment dans des régions éloignées du Québec, où les données d'observations se font rares.

Les objectifs de ce mémoire sont d'évaluer le potentiel des réanalyses à représenter le cycle hydrologique de surface au Québec, et leur performance en modélisation hydrologique sur des bassins versants du Québec. Dans un premier temps, une revue de littérature est établie pour mieux situer la recherche parmi les études précédemment réalisées dans ce domaine (Chapitre 1). Par la suite, le cycle hydrologique de surface et ses composantes provenant de quatre réanalyses récentes sont examinés sur le territoire du Québec (Chapitre 2). À la lumière des résultats apportés par cette étude comparative, les performances des quatre réanalyses en modélisation hydrologique sur des bassins versants du Québec sont évaluées.

(Chapitre 3). Finalement, une synthèse globale des résultats est établie, ainsi qu'une conclusion générale quant à l'applicabilité des réanalyses dans les études hydrologiques, au regard des résultats obtenus.

Le chapitre 2 ayant fait l'objet d'une soumission au Journal Of Hydrometeorology, le format de mémoire par article a été privilégié pour la rédaction de ces travaux. Par ailleurs, chacun des chapitres 2 et 3 présente sa méthodologie, ses résultats, ainsi que son analyse/discussion de façon indépendante, mais complémentaire.

CHAPITRE 1

REVUE DE LITTÉRATURE

Ce chapitre présente dans un premier temps les données de réanalyse, leur naissance et évolution, ainsi que les études publiées en lien avec cette recherche, c'est-à-dire les études comparatives des réanalyses, notamment les études de l'hydrologie de surface et du bilan d'eau. Dans un second temps, le concept de modélisation hydrologique est introduit, puis une synthèse des travaux publiés sur l'utilisation des données de réanalyse en modélisation hydrologique est présentée.

1.1 Les réanalyses

1.1.1 Qu'est-ce qu'une réanalyse?

À l'instar d'une prévision météorologique classique (utilisée par exemple dans un bulletin météo), une réanalyse est un jeu de données représentant l'état de l'atmosphère, ainsi que de sa surface sous-jacente. Le domaine spatial couvert par une réanalyse peut s'étendre sur la totalité du globe ou se focaliser sur une région particulière. Les variables de sortie sont représentées sur une grille régulière en 3 dimensions, et ce, sur une durée temporelle déterminée. Par exemple, les réanalyses récentes couvrent une période commençant en 1979 jusqu'au présent.

Les réanalyses sont générées par un modèle de prévision météorologique produisant une prévision à court terme, à un pas de temps (cycle de prévision) propre à chaque réanalyse. De plus, à chaque début de cycle de prévision, un processus d'assimilation de données d'observation permet de créer un nouvel état initial de l'atmosphère plus cohérent avec les observations pour le cycle suivant de prévision. Ce processus d'assimilation de données, commun aux réanalyses et aux prévisions météorologiques classiques, est appelé analyse. Le terme "réanalyse" vient de "retrospective analysis", car l'analyse est réalisée sur une période du passé. D'autre part, le modèle météorologique, la méthode d'analyse et tous les

algorithmes d'une réanalyse restent inchangés lors du processus de génération des données: ceci constitue la majeure différence entre les données de réanalyses et les produits de prévisions météo classiques. Les données d'observation assimilées peuvent provenir de différentes sources: satellites, bateaux, ballons sondes, avions, bouées ou stations météo entre autres. D'autre part, les variables fournies par les réanalyses sont nombreuses, et représentent de manière assez complète les processus physiques de l'atmosphère et de la surface: direction et intensité du vent, température, humidité, précipitation, évaporation, ruissellement, flux radiatifs, enneigement, etc.

1.1.2 Naissance et évolution des réanalyses

À partir des années 1970, le changement climatique commence à émerger en tant que secteur d'intérêt pour la recherche. En effet, certaines nations, les plus industrialisées principalement, se questionnent de plus en plus sur les impacts que peuvent avoir les activités humaines sur l'environnement, et sur les variations des cycles naturels climatiques préétablis. Ainsi, de nombreux outils et techniques scientifiques se perfectionnent, comme la modélisation météorologique, ou l'utilisation de statistiques poussées, dans le but de déceler de possibles tendances climatiques, via l'analyse des données du passé récent. Dans les années 80 et 90, les progrès considérables réalisés en modélisation météorologique et dans le développement des méthodes d'assimilation de données ont aussi eu comme effet d'accentuer la différence de qualité et la précision entre les données anciennement générées et les plus récentes. L'idée de procéder à des réanalyses météorologiques a alors émergé, afin de produire des données uniformes représentant l'atmosphère et la surface sur une période du passé récent (Bengtsson et Shukla, 1988).

Le Centre Européen de Prévision Météorologique à Moyen Terme (CEPMMT), ainsi que les National Centers for Environmental Prediction (NCEP) ont produit les premières réanalyses dans les années 90: respectivement, les réanalyses ERA-15 (Gibson, 1997) et du National Center for Atmospheric Research (NCEP-NCAR), aussi connue sous le nom R1 (Kalnay et al., 1996). Les faiblesses de ces premières réanalyses ont rapidement été mises en lumière, et

ont motivé la création d'une seconde génération: la réanalyse du Department Of Energy (NCEP-DOE), aussi connue sous le nom R2 (Kanamitsu et al., 2002) et la North American Regional Reanalysis (NARR), toutes deux développées par le centre NCEP (Mesinger et al., 2004), la réanalyse ERA-40 produite par le CEPMMT (Uppala et al., 2005) et la réanalyse japonaise JRA-25 (Onogi et al., 2007), développée par la Japan Meteorological Agency (JMA). L'étude des changements climatiques par les réanalyses nécessite de bien représenter les interactions complexes entre les océans, les continents et l'atmosphère à l'échelle globale. Or, ces interactions ne sont pas représentées dans les réanalyses de seconde génération. Pour cette raison, une troisième génération de réanalyses a vu le jour, notamment celles qui font l'objet de l'étude présentée ici: les réanalyses Climate Forecast System (CFSR) développée au centre NCEP (Saha et al., 2010), ERA-Interim du CEPMMT (Dee et al., 2011) et Modern-Era Retrospective analysis for Research and Applications (MERRA) produite par la NASA (Rienecker et al., 2011).

La conservation des résolutions spatiale et temporelle au cours du processus de génération des réanalyses représente un des principaux atouts de ces données vis-à-vis des données produites par un modèle de prévision météorologique, de même que le grand nombre de variables disponibles. Cependant, les réanalyses révèlent aussi des limitations à représenter correctement l'atmosphère. Ces limitations sont généralement liées à la phase critique du processus de génération des réanalyses: l'assimilation des données d'observation. En effet, dans les réanalyses, la densité des observations assimilées varie dans le temps et l'espace, ce qui peut entraîner l'apparition de fausses variabilités ou des tendances artificielles. Pour cette même raison, la fiabilité d'une variable peut varier fortement dans le temps ou l'espace. De plus, pour la plupart des réanalyses, l'assimilation de données, qui se fait de manière discrète dans le temps, force le modèle à se "plier" aux observations, et par conséquent le bilan d'eau ne se conserve pas. D'autre part, des différences significatives dans les variables peu contraintes peuvent être observées entre réanalyses, comme par exemple la précipitation ou les flux de surface. Différentes composantes du système climatique de la terre, comme les océans, les continents et l'atmosphère, peuvent être impliquées dans le processus de génération des réanalyses. Ainsi, contrairement aux anciennes réanalyses, les récentes CFSR,

ERA-Interim et MERRA intègrent de manière couplée ces différentes composantes pour mieux représenter les interactions climatiques globales entre ces composantes du système climatique. Cependant, dans le cas de ces réanalyses, malgré le fait que le processus de prévision est réalisé conjointement entre l'atmosphère et les surfaces océanique et continentale, les analyses (donc le processus d'assimilation) de ces composantes sont conduites indépendamment, ce qui peut entraîner l'apparition d'incohérences aux frontières à chaque début de cycle de prévision. C'est pourquoi ces réanalyses sont caractérisées de "semi-couplées".

Néanmoins, les réanalyses ont l'avantageuse particularité de pouvoir fournir de l'information climatique dans les régions où les mesures de données météorologiques sont rares, voire inexistantes. Par ailleurs, au Québec, la distribution spatiale des stations météorologiques est très irrégulière à travers la province. Elle présente une forte densité au sud, mais décroît rapidement vers le nord, jusqu'à l'absence presque totale de stations dans le grand nord. Les stations d'observation hydrométrique sont également rares dans ces régions, et les séries temporelles de débits observés sont souvent de courte durée et incomplètes (données manquantes). L'objectif de cette étude est donc d'évaluer la capacité des réanalyses à fournir des informations cohérentes sur le cycle hydrologique de surface au Québec, plus précisément des données de précipitation, évaporation et ruissellement fiables, d'abord à l'échelle de la province de Québec et plus particulièrement dans le nord.

1.1.3 Études comparatives des réanalyses

Depuis leur apparition, les réanalyses ont fait l'objet de nombreuses études scientifiques, dans le but de déceler leurs faiblesses et de promouvoir leurs atouts. La supériorité de la seconde génération de réanalyse par rapport à la première a été mise en évidence par Bosilovich et al. (2008) en ce qui concerne la représentation de la précipitation globale, pour les cas des réanalyses JRA-25 et ERA-40 versus R1 et R2. Dans la seconde génération, la réanalyse NARR tient une place particulière: elle assimile la précipitation contrairement aux autres réanalyses, s'étend sur un domaine « régional », soit l'Amérique du nord (Mesinger et al.,

2004), et possède par ailleurs une résolution spatiale horizontale fine de 32 km mieux adaptée aux études régionales. Bukovsky et Karoly (2007) ont examiné la précipitation de NARR. En se référant à un jeu de données d'observation sur grille, ils ont montré que NARR présente de meilleurs résultats par rapport aux réanalyses R2 et ERA-40 sur les États-Unis continentaux. Toutefois, un biais significatif de précipitation le long des frontières canado-américaines et dans l'Est du Canada apporte des questionnements sur la fiabilité de NARR. Les auteurs de cette étude ont donc invité les utilisateurs à prendre des précautions lors de l'utilisation de ces données dans ces régions. D'autre part, toujours sur les États-Unis continentaux, le modèle de surface de NARR ne semble pas distribuer annuellement les bonnes proportions d'eau précipitée en évaporation et ruissellement, en comparaison avec des données d'observation (Sheffield, Livneh et Wood, 2012).

CFSR est la plus récente réanalyse développée par le centre NCEP. Elle se démarque des autres réanalyses entre autres par le fait que son modèle de surface est alimenté directement par des précipitations observées, à la place de précipitation du modèle. Toutefois, ces précipitations modélisées ont tout de même fait l'objet d'analyse. En particulier, Wang et al. (2011) ont montré des améliorations dans la répartition temporelle des précipitations de CFSR sur les régions maritimes et océans, au regard d'autres réanalyses comme R1, R2 et ERA-40. Cependant, l'évaporation (disponible sous forme de flux de chaleur latente) semble plus importante sur tout le globe par rapport aux données d'observation utilisées comme référence. D'autre part, une brusque variation entre les années 1998 et 2001 a été mise en lumière concernant certaines variables, dont la précipitation. Zhang, Kumar et Wang (2012) ont par ailleurs fait de cet aspect le sujet de leur étude.

La réanalyse ERA-Interim du CEPMMT est une réanalyse globale récente, et se différencie de NARR, MERRA et CFSR par son système d'assimilation 4D-Var (four dimensional variational scheme). Ce schéma d'assimilation est plus complexe, et requiert une puissance de calcul substantiellement plus importante que le 3D-Var (de NARR, MERRA et CFSR), mais utilise néanmoins plus efficacement les données d'observation disponibles (Dee et al., 2011). L'hydrologie de surface de ERA-Interim a été analysée par Betts, Köhler et Zhang

(2009) sur les bassins versants de l'Amazonie, du Mississippi et Mackenzie. Les résultats montrent une amélioration générale de la précipitation et de l'évaporation d'ERA-Interim par rapport à son aînée ERA-40, vis-à-vis de données d'observation de référence.

La réanalyse MERRA, développée par la NASA, a, quant à elle, été conçue pour fermer le bilan hydrologique après chaque cycle d'analyse, par l'ajout d'un incrément qui aura comme objectif d'équilibrer le bilan d'eau à chaque itération (Rienecker et al., 2011). Ce cycle hydrologique a été examiné sur le globe entier par Bosilovich, Robertson et Chen (2011), dont les résultats ont montré une bonne fermeture du bilan hydrologique sur les surfaces continentales et océaniques. Néanmoins, un décalage dans l'équilibre du bilan d'eau est révélé à partir de 1999 (illustré par le changement de signe de l'incrément), probablement associé à l'introduction des données satellitaires Advanced Microwave Sounding Unit (AMSU).

Les réanalyses NARR, CFSR, ERA-Interim et MERRA ont aussi fait ensemble l'objet d'études de comparaison du cycle hydrologique aux échelles globale et régionale, sur des cycles annuel moyens et des moyennes à long terme (Lorenz et Kunstmann, 2012; Trenberth, Fasullo et Mackaro, 2011). Ces auteurs ont mis en valeur la similarité entre $P - E$ (précipitation moins évaporation) sur les terres et $E - P$ sur les océans, concernant les réanalyses ERA-Interim et MERRA. En d'autres termes, ces réanalyses montrent des bilans d'eau globalement plus équilibrés que NARR ou CFSR.

Par ailleurs, les travaux en cours de Essou et al. (2015) soumis au Journal of Hydrometeorology centrés sur l'étude de CFSR, MERRA, ERA-Interim et NARR sur le territoire des États-Unis pour la période 1979-2003, révèlent que la précipitation et les températures des réanalyses se comparent bien au jeu de données d'observations interpolées sur grille Santa Clara (Maurer et al., 2002). De plus, parmi les quatre réanalyses sur les États-Unis continentaux, NARR révèle globalement les précipitations les plus fidèles aux données sur grille de Santa Clara: biais moyens saisonniers relativement faibles ($\pm 10\%$), variabilité semblable et corrélation supérieure à 0,8 vis-à-vis des données Santa Clara.

1.2 La modélisation hydrologique

Les réanalyses peuvent apporter des informations climatiques dans les régions où les mesures de données météorologiques sont quasi-inexistantes, ce qui est notamment le cas du nord du Québec. Toutefois, est-ce que les réanalyses sont en mesure de produire des séries de débits fiables simulés par l'intermédiaire d'un modèle hydrologique, notamment dans les régions sans donnée d'observation disponible? Avant d'y répondre, les thématiques de recherche relatives à cette question sont introduites, plus particulièrement la modélisation hydrologique, et l'historique des études récentes associant les réanalyses à la modélisation hydrologique.

1.2.1 Qu'est-ce qu'un modèle hydrologique?

Un modèle hydrologique est un ensemble d'équations mathématiques qui permet, à partir des données météorologiques d'un bassin versant (typiquement la précipitation et la température), de simuler les débits de la rivière à l'exutoire de ce bassin. En hydrologie, les modèles les plus courants sont les modèles déterministes à base physique ou conceptuels. Un modèle à base physique nécessite la connaissance des caractéristiques propres d'un bassin versant pour la résolution des équations de la conservation de l'eau et de l'énergie, comme les propriétés des couches de sol, la couverture du sol (végétation, forêt, zones urbaines, etc.), ou encore le relief. Cependant, les propriétés intrinsèques d'un bassin versant sont souvent difficiles à mesurer, c'est pourquoi le choix du modèle conceptuel est souvent favorisé.

Un modèle conceptuel représente les processus qui façonnent le cycle de l'eau sur un bassin versant sous forme de concepts (réservoirs et relations décrivant les écoulements verticaux entre ceux-ci ainsi que l'écoulement horizontal qui en est issu, par exemple.). Ces concepts ne représentent pas exactement la réalité, mais tentent tout de même de s'en rapprocher (Pechlivanidis et al., 2011). En effet, certains aspects d'un bassin versant, par exemple une couche de sol, peuvent facilement être identifiés comme un réservoir d'eau dont les caractéristiques sont ajustables. Les grandeurs ajustables d'un modèle sont appelées paramètres, et les valeurs de ces paramètres sont généralement propres à chaque bassin versant. Cependant, contrairement à l'exemple de la couche de sol, les paramètres d'un

modèle n'ont pas nécessairement de sens physique. Dans plusieurs études, telle que celle qui est menée dans le présent mémoire, l'objectif d'un modèle est de simuler, de la manière la plus réaliste possible, le débit à l'exutoire d'un bassin versant. Pour ce faire, il est nécessaire de calibrer le modèle, c'est-à-dire d'ajuster les valeurs des paramètres du modèle. Dans le cas d'un bassin versant jaugé, c'est-à-dire, un bassin versant dont le débit à l'exutoire a été enregistré au cours du temps, la série de débits observés est comparée à la série de débits simulés via un critère d'optimisation, aussi appelé fonction objectif. Calibrer un modèle revient donc à obtenir des valeurs de fonction objectif satisfaisantes, telles que la série de débits simulés reproduise la série de débits observés de la manière la plus réaliste, en faisant varier par itération les valeurs des paramètres du modèle. Dans le cas des bassins versants non jaugés, ce processus d'optimisation n'est pas envisageable puisque les séries de débits observées sur ces bassins ne sont pas disponibles. Il existe cependant des techniques permettant de contourner cette contrainte, mais elles ne seront néanmoins pas détaillées ici, car ne faisant pas l'objet de ce mémoire. L'optimisation d'un modèle conceptuel peut s'effectuer manuellement, toutefois l'utilisation d'un algorithme d'optimisation est habituellement favorisée. Le calage manuel peut s'avérer laborieux, en particulier lorsque le nombre de paramètres à caler est élevé. Il est subjectif mais, en contrepartie, tire profit de l'expertise de l'hydrologue. Le calage automatisé est entièrement objectif et reproductible, et implique essentiellement du temps « machine » pour la réalisation des calculs itératifs. Par ailleurs, lorsque ces paramètres sont nombreux, plusieurs combinaisons de paramètres peuvent correspondre à des valeurs de fonctions objectif équivalentes. Cet aspect connu de l'étape de calage d'un modèle conceptuel est aussi appelée équi-finalité des paramètres (Beven, 2006).

Les modèles hydrologiques peuvent aussi se différencier selon leurs approches spatiales. Deux grands types coexistent à ce jour, les modèles globaux et les modèles distribués (Moradkhani et Sorooshian, 2008). Le modèle global considère les grandeurs physiques comme uniformes en tout point du bassin versant, par exemple une même hauteur de précipitation, une même température et une même évaporation sur tout le bassin. Le modèle distribué est quant à lui constitué de plusieurs entités différentes, pouvant être considérées

comme des sous-bassins. Chaque sous-bassin est homogène, mais ceux-ci considèrent des grandeurs physiques distinctes entre eux (précipitation, évaporation, etc.). Par ailleurs, les modèles distribués sont généralement des modèles à base physique. Ces modèles peuvent apparaître comme étant plus proches de la réalité que les modèles globaux, cependant ils nécessitent généralement beaucoup plus de puissance de calcul que les modèles globaux et leur mise en place sur un bassin nécessite beaucoup plus de travail, associé notamment au traitement des données physiographiques.

Les applications de la modélisation hydrologique sont nombreuses. En effet, cette dernière s'avère très utile dans le cadre de la gestion des ressources hydriques disponibles, à des fins agricoles ou industrielles. Elles sont aussi essentielles lors de la phase de dimensionnement d'ouvrages d'ingénierie tels qu'un barrage, un pont, une route ou encore une conduite. La modélisation joue aussi une part cruciale dans le domaine de la prévision opérationnelle, par exemple la prévision à court terme des débits extrêmes en rivières et des potentielles inondations, mais aussi dans le cadre de la projection climatique, c'est-à-dire, dans les estimations à long termes de l'évolution des régimes hydriques, potentiellement perturbés par les changements climatiques.

1.2.2 Modélisation hydrologique et réanalyses

La modélisation hydrologique peut aussi être utilisée comme un outil de validation des données météorologiques, dont les réanalyses font partie (Hagemann et Gates, 2001; Maurer et al., 2001). Inversement, lorsque qu'une région d'intérêt contient très peu ou pas de données d'observation mesurées à l'aide de stations météorologiques, les réanalyses sont en mesure de fournir ces informations. Ainsi, différentes études ont été menées sur l'analyse de la performance des réanalyses en modélisation hydrologique, notamment avec les réanalyses CFSR (Dile et Srinivasan, 2014; Fuka et al., 2014; Lauri, Räsänen et Kummu, 2014), NARR (Choi et al., 2009; Eum et al., 2014) et ERA-Interim (Lauri, Räsänen et Kummu, 2014).

Dile et Srinivasan (2014) ont comparé les résultats de simulations hydrologiques basées sur CFSR et sur des stations d'observations du bassin versant du Nil Bleu, en Éthiopie, et ont indiqué que globalement, les simulations basées sur les mesures des stations produisent des débits plus fidèles aux débits observés que celles basées sur CFSR. Néanmoins, les simulations basées sur la réanalyse fournissent tout de même des résultats exploitables dans les régions pauvres en données d'observation météorologiques.

Par ailleurs, Fuka et al. (2014) ont modélisé des écoulements en rivière sur cinq bassins versants représentant des régimes hydro-climatiques différents (aux USA et en Éthiopie), en utilisant les données météorologiques de CFSR. Ils ont conclu que cette réanalyse a la capacité de fournir des séries de débits simulés fiables en modélisation hydrologique sous différents régimes hydro-climatiques, et ajoutent que les résultats de modélisation avec CFSR peuvent atteindre des valeurs similaires, voir être supérieurs aux résultats de modélisation avec les données de stations d'observation. Selon les auteurs, ceci serait probablement lié au fait que les données de CFSR sont moyennées sur des surfaces (environ 1000 km²) comparables à la superficie de certains bassins (bassin de Gumera: 1200 km²), ou du moins, sont plus représentatives de la surface des bassins versants que ne le sont les données de stations.

Choi et al. (2009) ont examiné l'applicabilité des données de température et de précipitation de NARR pour la modélisation hydrologique, sur des bassins versants du Manitoba au Canada, et ont montré que les résultats de calages du modèle hydrologique SLURP (Semi-distributed Land Use-based Runoff Processes) avec NARR ont été légèrement inférieurs aux résultats de calage avec des données d'observations. Ils en ont conclu que ce modèle hydrologique était en mesure d'être étalonné correctement avec les données NARR, et de fournir des informations hydrologiques fiables, notamment dans les régions où les stations météorologiques se font rares. De plus, les auteurs ont noté que lorsque le modèle calibré avec les données d'observations était utilisé avec les données NARR, le ruissellement simulé présentait un biais systématique de 20 % par rapport à la série observée. Malgré les biais que peuvent contenir les données de réanalyses, si la physique des phénomènes

climatiques est bien représentée au sein de la précipitation ou de la température par exemple, le modèle hydrologique est capable de compenser le manque ou le surplus d'eau via l'ajustement de ses paramètres.

De même, Eum et al. (2014) ont examiné et comparé trois jeux de température et précipitation au Canada, incluant NARR, l'analyse canadienne des précipitations CaPA (Fortin, Roy et Donaldson, 2014) et les données sur grilles NRCAN (voir section 2.3.2) et ont aussi effectué des comparaisons de performance avec le modèle hydrologique VIC (Variable Infiltration Capacity) sur le bassin versant de l'Athabasca (réparti sur l'Alberta et le Saskatchewan au Canada). Les auteurs ont observé des différences systématiques entre les jeux de données. De même, les résultats de simulations hydrologiques ont montré de grandes différences dans les débits simulés avec les trois jeux de données, particulièrement dans les zones montagneuses. Par ailleurs, notons que le modèle VIC n'a pas été recalibré avec chaque jeu de données testé, le jeu de paramètres utilisé demeure donc le même pour chaque simulation.

D'autre part, Lauri, Räsänen et Kummu (2014) ont examiné la performance en modélisation hydrologique de CFSR et ERA-Interim sur le bassin versant tropical du fleuve Mekong (Asie du Sud-Est), et montré que le couplage entre précipitations observées et températures des réanalyses produit des performances hydrologiques analogues aux simulations réalisées exclusivement avec des observations. Les auteurs ont ajouté que les réanalyses CFSR et ERA-Interim présentent des problèmes de distribution spatiale de la précipitation sur ce bassin versant, et par conséquent, n'ont pas généré de simulation hydrologique satisfaisante.

Les travaux de Essou et al. (2015) précédemment cités (section 1.1.3) ont aussi porté sur l'évaluation des réanalyses CFSR, MERRA, ERA-Interim et NARR en modélisation hydrologique sur des bassins versants des États-Unis, par région climatique. Ainsi, en comparant les débits simulés par les réanalyses via le critère de Nash-Sutcliffe avec les séries de débits observés de la base de données MOPEX (Model Parameter Estimation Experiment, (Duan et al., 2006)) , ils ont montré que les performances des réanalyses se comparent bien à

celles des observations, excepté dans les régions humides continentale et subtropicale pour ERA-I, CFSR et MERRA Essou et al. (2015).

Les études hydrologiques au moyen des réanalyses CFSR, MERRA, ERA-Interim et NARR sont relativement nombreuses. Cependant, ces études se positionnent généralement d'un point de vue global, ou sur des grands bassins versants. Par ailleurs, les évaluations de la performance de ces réanalyses en modélisation hydrologique restent relativement peu nombreuses à ce jour, et même inexistante sur la province de Québec. De plus, contrairement aux travaux présentés ici, peu de ces études s'intéressent à plus de deux réanalyses. Ainsi, l'étude du cycle hydrologique de surface et l'évaluation de la performance en modélisation hydrologique des réanalyses CFSR, MERRA, ERA-Interim et NARR sur le Québec et ses bassins versants pourraient apporter beaucoup d'information concernant l'applicabilité de ces données de réanalyses à des études hydrologiques, notamment dans les régions où les mesures d'observation sont peu ou pas disponibles.

CHAPITRE 2

USE OF REANALYSIS DATA TO ASSESS THE TERRESTRIAL BRANCH OF THE WATER CYCLE: A CASE STUDY OF FOUR DATASETS OVER QUEBEC, CANADA

Florent Sabarly¹, Gilles Essou¹, Philippe Lucas-Picher^{1,2}, Annie Poulin¹ et François Brissette¹

¹Département de Génie de la Construction, École de Technologie Supérieure,
1100 Rue Notre Dame Ouest, Montréal, Québec, Canada, H3C 1K3

²Département des sciences de la Terre et de l'atmosphère, Université du Québec à Montréal,
201 Avenue du Président Kennedy, Montréal, Québec, Canada H2X 3Y7

Article soumis à « Journal of Hydrometeorology » en Juin 2015

2.1 Abstract

Reanalyses have the potential to provide meteorological information in areas where few or no traditional observation records are available. The terrestrial branch of the water cycle of the reanalyses CFSR, MERRA, ERA-Interim and NARR, is examined over Quebec, Canada. Precipitation, evaporation, runoff and water balance are studied using observed precipitation and streamflows, according to three spatial scales: (1) the entire province of Quebec, (2) five regions derived from a climate classification, and (3) eleven river basins. The results reveal that MERRA provides a relatively closed water balance, while a significant residual was found for of the other three reanalyses. MERRA and ERA-Interim seem to provide the most reliable precipitation over the province. On the other hand, precipitation from CFSR and NARR do not appear to be particularly reliable, especially over southern Quebec, as they almost systematically showed the highest and the lowest values, respectively. Moreover, the partitioning of precipitation into evaporation and runoff from MERRA and NARR does not agree with what was expected, particularly over southern, central and eastern Quebec. Despite the weaknesses identified, the ability of reanalyses to reproduce the terrestrial water cycle of the recent past (i.e. 1979 - 2008) remains globally acceptable. Nonetheless, their

potential to provide reliable information must be validated by comparing reanalyses directly with weather stations, especially in remote areas.

2.2 Introduction

Since the end of the last century, interest in the study of climate change has grown considerably. Looking at observations or data produced by global climate models representing the recent past, and analyzing them using statistics could be relevant to detecting possible trends regarding climate change. In the '80s and '90s, considerable improvements in weather forecasting and numerous upgrades of model and data assimilation methods contributed to the notion of reanalysing the recent past (Bengtsson et Shukla, 1988). Briefly, a reanalysis aims to provide the best estimation of the state of the atmosphere, ocean and land surface from the recent past. The term reanalysis stands for Retroactive Analysis, since the analysis is done on a past period extending up to the near present. Reanalyses are three-dimensional gridded datasets produced by a weather forecasting model. Two main characteristics can define these datasets. Firstly, a data assimilation scheme is used to integrate observations from different sources in order to provide the most coherent state of the atmosphere. Secondly, the assimilation scheme and the forecasting model of a reanalysis remain unchanged during the entire simulation period. As such, inconsistencies that might be induced by continuous updates of the data production system are avoided.

Many observations measured from different sources, such as radio sounding, aircrafts, boats, satellites, surface sensors, buoys, etc., are assimilated in the production of a reanalysis. Moreover, the large range of available reanalysis data products provides information such as radiative fluxes, wind, temperature, humidity, precipitation, albedo, snow, vegetation and land cover, to name just a few. The global coverage and the huge range of available variables with a consistent time and space resolution during the simulated period represent some of the main benefits that reanalyses provide to climate studies, including a smaller bias and a finer spatial resolution from one generation of reanalyses to the next. Nevertheless, reanalyses also show some limitations. Indeed, the reliability of some variables may significantly vary in

time and space. Moreover, the evolution of the number and quality of assimilated observations may introduce some artificial variability and trends. As well, the water balance is rarely conserved, and reanalyses sometimes show substantial differences between variables, such as precipitation, that are not directly constrained by assimilation.

In the '90s, the European Centre for Medium-Range Weather Forecasts (ECMWF) and the National Centers for Environmental Prediction (NCEP) began producing the first global reanalyses, the ECMWF ERA-15 reanalysis (Gibson, 1997) and the NCEP National Center for Atmospheric Research Reanalysis (NCEP-NCAR), also known as R1 (Kalnay et al., 1996). Armed with awareness about the limitations of these first two datasets, a second generation of reanalyses was produced, namely, the Department Of Energy reanalysis (NCEP-DOE), also known as R2 (Kanamitsu et al., 2002) and the North American Regional Reanalysis (NARR) (Mesinger et al., 2004), both developed by NCEP, the ECMWF ERA-40 reanalysis (Uppala et al., 2005) and the Japanese 25-yr. reanalysis JRA-25 (Onogi et al., 2007) developed by the Japan Meteorological Agency (JMA). Recently, a third generation of reanalysis was developed, including the Climate Forecast System Reanalysis (CFSR) (Saha et al., 2010) from NCEP, the ECMWF ERA-Interim reanalysis (Dee et al., 2011) and the Modern-Era Retrospective analysis for Research and Applications (MERRA) produced by NASA (Rienecker et al., 2011).

In addition to their application to climate studies, reanalyses have the potential to provide climate information in areas that are sparsely inhabited or with limited surface observations (in space or time), for applications such as water resource management and hydrological modelling. For the latter, examining precipitation, evaporation, runoff and the water balance for the terrestrial branch of the water cycle obtained through reanalyses should provide useful information.

Bukovsky et Karoly (2007) compared the precipitation of NARR, R2 and ERA-40 using a set of gridded observations. Exploring the spatial distribution and the diurnal and annual cycles, they revealed that NARR showed better results than the other two reanalyses over the

continental United States. Nevertheless, these authors recommended that users should proceed cautiously when looking at the rest of the North American domain, especially along the US borders and in South-eastern Canada, where the overall NARR precipitation is strongly underestimated. Among all existing reanalyses, NARR is the only one that assimilates precipitation, and in which the quality of the simulated precipitation strongly depends on the quality of observed data and on the assimilation process (Mesinger et al., 2004). Zhang, Kumar et Wang (2012) focused on the change in the global average of precipitation from CFSR during the 1998-2001 period. They demonstrated that an interaction between the bias of the data assimilation model and the nonstationarity in the ingestion of some observed data is the source of the global average increase in the CFSR precipitation after 1998. Bosilovich, Robertson et Chen (2011) evaluated the water balance of MERRA over the entire globe. Since this reanalysis was configured to include in the water balance the residual generated by the data assimilation process, the water balance over the land and oceans is closed. However, a notable shift in annual water balance was identified, starting in 1999, which coincides with the beginning of Advanced Microwave Sounding Unit (AMSU) radiance assimilation. Sheffield, Livneh et Wood (2012) identified noteworthy differences in NARR evaporation and runoff, compared to two offline land surface model simulations (Noah v2.7.1 and Variable Infiltration Capacity - VIC), using observational runoff estimates over the continental United States. NARR (which uses a previous version of Noah) and Noah simulations present an overestimation of annual evaporation, and runoff ratios (simulated runoff divided by observed runoff) that are 50% lower than the VIC simulation results. Regarding NARR, the authors identified these differences as being mainly related to the evaporation component of the Noah model, versus other factors such as atmospheric forcings or biases induced by precipitation assimilation into NARR. Lorenz et Kunstmann (2012) investigated the closure of the water balance (i.e. balance between precipitation, evaporation, surface runoff and moisture flux) of ERA-Interim, MERRA and CFSR over the entire globe, and found that ERA-Interim is the reanalysis that likely provides the most reliable rainfall estimates globally. Moreover, they showed that the long-term mean $P - E$ (precipitation minus evaporation) values of MERRA and ERA-Interim over land are similar to the $E - P$ values over the oceans, confirming that these two reanalyses show better global water

balance conservation than CFSR. Furthermore, the change in the number of assimilated data in CFSR around 1998 revealed by Zhang, Kumar et Wang (2012) leads to a substantial imbalance between $P - E$ over the land and oceans.

The province of Quebec in Canada has abundant freshwater resources. Its numerous lakes and rivers play a fundamental role in local wildlife and flora sustainability. On the other hand, there is great interest in managing water resources, particularly in terms of hydroelectricity production. Quebec has a surface area greater than 1,5 million km², and is characterized by five different climate regimes (Bukovsky, 2011). However, the spatial distribution of meteorological stations varies across the province, dense in the south, decreasing to the north, to almost non-existent in the far north. Therefore, reanalyses should be useful in providing substantial information, particularly in these remote regions with little observational data. This study focuses on the assessment of the components of the terrestrial branch of the water cycle (Peixoto et Oort, 1992) of four recent reanalyses, CFSR, ERA-Interim, MERRA and NARR, over the province of Quebec, and especially on the assessment of the reliability of the four reanalyses in representing the hydrological cycle components in the northern regions of the province. The analysis is divided into three parts. In the first part, the long-term mean water balance is examined over the entire territory of Quebec. Secondly, the mean annual cycle of the terrestrial water cycle components and annual water balance are analysed according to the climate classification of Bukovsky. Thirdly, the precipitation, runoff and E/P ratio (evaporation divided by precipitation) are examined over river basins representing the different hydrological regimes of the province. Section 2.3 details the methodology and the main characteristics of the datasets used in this study. Section 2.4 presents the results and section 2.5 provides the concluding remarks.

2.3 Data and methods

The reanalysis and observational datasets that were used in this study are first described, and then the methodology follows.

2.3.1 Reanalysis datasets

The widely used NARR, CFSR, MERRA and ERA-Interim reanalyses were chosen to evaluate the terrestrial branch of the water cycle over the province of Quebec. These reanalyses benefited from improvements obtained from the preceding generation of reanalyses, and thus represent the most advanced and suitable products available for carrying out this study. This section describes the relevant properties of each reanalysis as used in this study, including their similarities and differences. For the analysis, we selected the 1979-2008 period since the 30-year period is widely used in climate studies. Moreover, the selected period starts in 1979, as it is the first year covered by the four selected reanalyses. All the references and main properties of the datasets used in this study are summarized in Tableau 2.1.

Tableau 2.1 Main properties summary of reanalysis datasets

	CFSR	ERA-Interim	MERRA	NARR
Organization	NCEP	ECMWF	NASA	NCEP
Period	1979-present	1979-present	1979-present	1979-present
Horizontal resolution	0.3° (~33 km)	0.75° (~83 km)	2/3° x 1/2° (~74 x 56 km ²)	~32 km
Assimilation approach	3D-VAR	4D-VAR	3D-VAR	3D-VAR
Particularity	Surface model forced by observed <i>P</i>	Uses 4D-VAR scheme	Designed to balance the water cycle	Assimilates precipitation
References	Saha et al. (2010)	Dee et al. (2011)	Rienecker et al. (2011)	Mesinger et al. (2004)

The recent global ECMWF ERA-Interim reanalysis includes an assimilation system based on a 4D-VAR approach, which is more complex and computationally intensive than the 3D-VAR used by CFSR, NARR and MERRA, but uses more efficiently available observations (Dee et al., 2011). Most of the surface fields are available every three hours and every six hours for the atmospheric fields, on a 0,75-degree regular grid (about 83-km).

NARR, which covers the North American continent and parts of the North Atlantic and Pacific Oceans, has the particularity of assimilating observed precipitation, contrary to the other three reanalyses. In fact, before being assimilated, the observed precipitation in NARR is converted into latent heat. Additional details about how NARR is generated and about precipitation assimilation can be found in Mesinger et al. (2004). NARR outputs are available every three hours, with a 32-km horizontal resolution (about 0,3-degree).

CFSR is the first coupled global atmosphere, ocean, land surface and sea ice system reanalysis (Saha et al., 2010). As for ERA-Interim and MERRA, the atmosphere, ocean, land surface and sea ice models share boundary data during the forecasting process. However, the analysis of each model component in CFSR, ERA-Interim and MERRA is processed separately. Both CFSR and NARR use the Noah land surface model, but in CFSR, the Noah model has the particularity to be forced with observed precipitation, instead of using the precipitation generated by the atmospheric model, which is considered too biased (Saha et al., 2010). Furthermore, Meng et al. (2012) explained that « previous studies have shown nontrivial biases in the Global Data assimilation System precipitation (Gottschalck et al., 2005). Such a bias over land often leads to biases in many simulated land surface variables ». Surface fields of CFSR are available at a resolution of 0,3-degrees (about 32-km), and every six hours.

MERRA is a global reanalysis simulated on a 2/3-degree longitude and 1/2-degree latitude regular grid, with hourly outputs. The assimilation process of MERRA is similar to that of CFSR, which is the Gridpoint Statistical Interpolation scheme, developed at NCEP. In addition, the main specificity of MERRA consists in the use of an incremental analysis update (IAU) procedure to improve water balance conservation (for further explanations, see Rienecker et al. (2011)).

2.3.2 Observational datasets

This section describes the NRCAN observationally-based gridded precipitation, as well as the (cQ)² observed streamflow time series used as reference in this study. The NRCAN dataset provides daily Canada-wide precipitation, minimum and maximum temperatures on a 10-km resolution regular grid (Hutchinson et al., 2009). It was originally produced to support studies requiring daily data, for instance, hydrological modelling, agricultural and forestry applications, extreme event analysis, etc. The gridded precipitation was obtained from the interpolation of measurements at individual stations, whose number varies in time from about 2000 to 3000 for the 1961-2010 period. The ANUSPLIN model was applied for the interpolation. It uses a trivariate thin plate smoothing splines method to model the spatial distribution of precipitation as functions of latitude, longitude and elevation, across Canada. Figure 2.1 shows the locations of the 956 weather stations available to generate the daily precipitation over the province of Quebec, provided by Environment Canada.

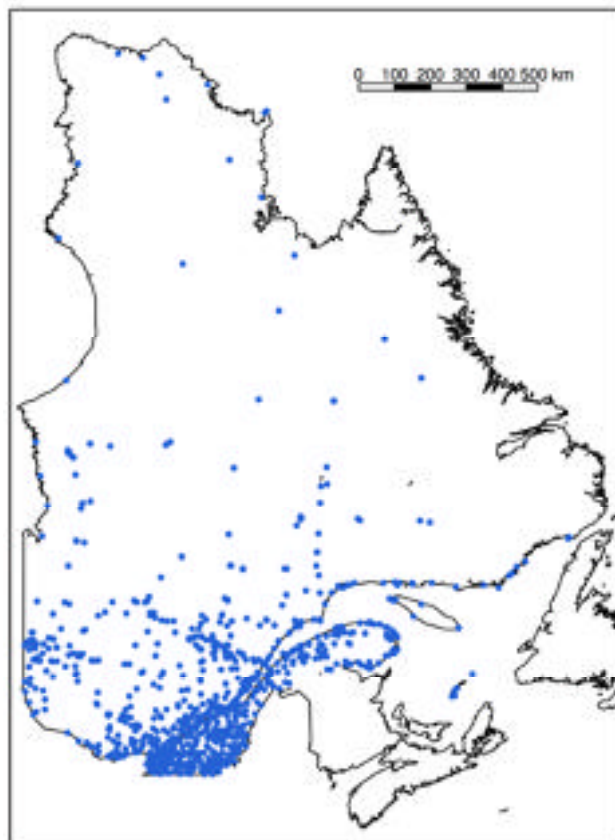


Figure 2.1 Location of the 956 Environment Canada weather stations used to generate the daily NRCAN dataset (Hutchinson et al., 2009). The stations are represented only over the eastern territories of Canada

The (cQ)² database (Impact des Changements Climatiques sur l'hydrologie(Q) au Québec) contains daily outlet streamflows, names and surface areas, as well as center and contour coordinates of 306 river basins over the province of Quebec (Arsenault et Brissette, 2014). This database was jointly produced by Hydro-Quebec, Rio Tinto Alcan and the Centre d'Expertise Hydrique du Québec, in order to unify water resource information and simplify access to hydrometric data.

Both observational datasets have some known limitations mainly related to streamflow measurements in (cQ)² and to decreasing spatial density of weather stations used to produce the NRCAN dataset from southern to northern Quebec. For instance, over the Saguenay-Lac Saint Jean and west of the Côte Nord regions, NRCAN shows inconsistent precipitation between 1995 and 1996, and significant underestimation from 2004 to 2008, due to the malfunction of a weather station (C. Guay 2015, personal communication). Moreover, measuring consistent streamflow time series in northern remote areas is challenging and may introduce some errors. Therefore, some of the (cQ)² time series have been post-processed to correct some inconsistencies; for instance, natural streamflow reconstruction over regulated river basins or measurement correction due to inaccurate measurements of river flows under ice cover. Despite the application of post-processing techniques, some biases remain, and streamflow time series in the southern part of the province may be more reliable than those in the northern regions. Moreover, as the density of observed data integrated into the NRCAN decreases from South to North, the reliability of this dataset may be more questionable over northern regions.

2.3.3 Methodology

This paper examines the terrestrial water cycle at three different regional scales (entire Province of Quebec, climatic regions within the province, and river basins). The terrestrial water balance of the reanalyses can be written following Equation (2.1):

$$\frac{\partial W}{\partial t} = P - E - R - RES \quad (2.1)$$

where W is the surface water storage (mm), P is the precipitation (mm day⁻¹), E is the total surface evaporation (mm day⁻¹) and R is the total runoff (mm day⁻¹), including the subsurface runoff. The term RES (mm day⁻¹) is added in Equation (2.1) to close the reanalysis water balance. Indeed, during the assimilation process of reanalysis, a new state of the atmosphere is generated, different from the one simulated by the model, but closer to observations. Atmospheric states are then discontinuous, which unbalances the water balance. In other words, the RES term may also be viewed as an estimate of the overall error in the water balance (Roads et al. 2003). In this study, Equation (2.1) is rearranged into Equation (2.2):

$$B = \frac{P - E - R}{P} \times 100 \quad (2.2)$$

where B is the relative water balance (% of P), which includes both RES and the change of surface water storage. Considering climatic time scales, the temporal change in surface water storage (soil water and snow water equivalent) is assumed to be negligible (Kleidon et Schymanski, 2008). In this case, B equals RES . This assumption is not completely verified on an annual time scale. Nevertheless, Roads et al. (1998) concluded that the RES term should be the most important contribution in comparison with the change in water storage.

Precipitation, evaporation and runoff from CFSR, ERA-Interim, MERRA and NARR were first downloaded from their respective websites for the period going from 1979 to 2008 (30 years). Each variable is aggregated on a daily basis, and computed in millimeters per day.

The evaporation includes evaporation from the bare soil, transpiration from vegetation, interception loss and snow sublimation. For CFSR, the evaporation component is derived from the latent heat flux (in W m^{-2}), as the evaporation is not directly available. Equation (2.3) was used to compute the evaporation:

$$E = \frac{H_f}{L_e(T)} \times C \quad (2.3)$$

where E is the evaporation (mm day^{-1}), H_f the latent heat flux (W m^{-2}), $L_e(T)$ is the latent heat of vaporization (J kg^{-1}), and C is a constant to convert the evaporation to a daily time step ($C=86400 \text{ s}$). Although the latent heat of evaporation depends on the surface temperature T , its influence on the evaporation can be neglected (Lorenz et Kunstmann, 2012). Therefore, the latent heat of evaporation is approximated to $2,5 \times 10^6 \text{ J kg}^{-1}$ and Equation (2.3) can be simplified as follows (Equation (2.4)):

$$E = H_f \times 0.03456 \quad (2.4)$$

In the case of the NARR and MERRA datasets, baseflow and surface runoffs, which are provided separately, are summed to compute the total runoff. The analysis that was carried out in this paper is divided into three parts, according to the three different spatial scales that are considered. Long-term and annual time scales have also been computed, as introduced hereafter.

2.3.3.1 Long-term mean of the hydrological cycle components over the province of Quebec

In the first part of this study, precipitation P , evaporation E , runoff R , and the relative water balance B from the four reanalyses are averaged at a daily time scale over the period 1979 - 2008 for each reanalysis tile. Values of P , E and R are computed on each tile following Equation (2.5):

$$X = \sum_{i=1}^n \frac{x_i}{n} \quad (2.5)$$

where X (mm day^{-1}) is the averaged variable P , E or R , x_i the value of P , E or R for day i and n the total number of days in the period 1979 - 2008. The relative water balance B is computed from the averaged values of P , E and R using Equation (2.2). To consider only data over the land, land-sea masks of the four reanalyses are applied. As MERRA natively provides land cover fractions, a threshold of 0.5 was used to distinguish the land from the water surfaces.

2.3.3.2 Hydrological cycle components over the climatic regions

In the second part of this study, spatial averages of P , E , R and B are computed over subregions of Quebec, according to the climate classification of Bukovsky (2011). This classification was created to provide a consistent climate division, as part of the North American Regional Climate Change Assessment Program (NARCCAP). The climate classification is based on a simplification of ecoregions of Ricketts et al. (1999). According to this classification, Quebec is divided into five climatic regions: Great Lakes (GL), North Atlantic (NA), East Boreal (EB), East Taïga (ETA) and East Tundra (ETU). Figure 2.2 (a) shows the five climatic regions over the province of Quebec, at the resolution of the CFSR dataset. Mean annual cycle of P , E and R are produced (one value per month per variable) following Equation (2.6):

$$Y = \sum_{j=1}^k \sum_{i=1}^p \frac{x_{ij}}{p \times k} \quad (2.6)$$

where x_{ij} is the value of P , E or R (mm day^{-1}) for day i and tile j , p is the number of reanalysis tiles within a climatic region, k is the total number of days of a month from every years of the period 1979 - 2008 (for example, $k=930$ for January). Finally, one value (Y , in mm day^{-1}) is computed for each month. The NRCAN precipitation is used as the reference for precipitation, and computed similarly to the reanalyses precipitation.

The relative water balance B is averaged annually (one value per year) for each climatic region, from October 1st to September 30th of the following year. As such, the snowpack is supposed to completely accumulate and melt, and then to be transferred to runoff during the year. Therefore, there are 29 annual values, since the periods of January 1st to September 30th of 1979, and of October 1st to December 31th of 2008 are not considered. The annually averaged value of P , E and R are first computed using Equation (2.6), except that k is now the number of days within the period from October 1st to September 30th of the following year. The relative water balance is then calculated with Equation (2.2) for each year.

2.3.3.3 Hydrological cycle components within river basins

In the third part of this study, total P , R and the E/P ratio (evaporation divided by precipitation) are computed annually, from October 1st to September 30th of the following year, over 11 selected river basins from the (cQ)² database. To further investigate precipitation from the four reanalyses, the temporal correlation coefficient between the daily time series of precipitation from each reanalysis and from NRCAN is computed over the river basins of interest. The river basins were selected according to three main criteria: (1) their representativeness of the six major hydrologic regimes of Quebec, namely South, Center, North, Gaspésie, Côte Nord and Arctic; since the (cQ)² streamflow time series contain missing data, especially during the winter low flow period, (2) the basins with the longest period without any missing data, and (3) the longest temporal coverage were favoured. Rivers in Arctic and North regions flow to the west and north, whereas those in Côte Nord, Gaspésie, South and Center flow into the St. Lawrence River. These regimes correspond fairly well with the Bukovsky climatic regions, with a distinction being made between the Center and Côte Nord, as compared to the unique East Boreal climate region. Each hydrological regime is represented by two river basins, except for the Arctic, where only one river basin showed consistent observed streamflows. Furthermore, one of the two river basins in the South is formed by three small aggregated river basins; their streamflow time series were spatially aggregated, while their surface areas were summed up. This way,

the number of reanalysis gridpoints within the river basin increases, providing more relevant information.

Figure 2.2 (b) shows the hydrologic regions and the selected river basins of $(cQ)^2$ of Quebec. Among the 306 river basins, some are subbasins, and are not represented in Figure 2.2 (b). The calculation method is the same as over the climatic regions, except that the daily values are spatially averaged over river basins instead of climatic regions. Moreover, averages are computed as follows: (1) when at least four dataset gridpoints are located inside the contour of a river basin (or aggregated river basin), a simple arithmetic mean is used; (2) otherwise, the Thiessen's polygons method is applied (Rhynsbarger, 1973), attributing specific weights to the four closest gridpoints inside or outside the river basin(s) contour. Regarding the runoff analysis, a runoff estimate is calculated using the $(cQ)^2$ streamflows relatively to the surface area of each of the 11 river basins. Moreover, NRCAN precipitation is used as reference data, as well as an estimated E/P ratio derived from NRCAN precipitation and $(cQ)^2$ streamflows. The reference ratio is calculated by total annual streamflow volumes ($m^3 s^{-1}$) being converted into runoff ($mm day^{-1}$) according to the surface area of each river basin. Assuming that the water balance of observed data is closed, an estimation of evaporation is computed using the Equation (2.7):

$$E = P - R \quad (2.7)$$

where E is the evaporation estimate derived from observations ($mm day^{-1}$), P is the NRCAN precipitation ($mm day^{-1}$) and R is the runoff ($mm day^{-1}$) derived from observed streamflows. As for the preceding calculations, Equation (2.7) is used on an annual basis from October 1st to September 30th of the following year. The E/P ratio is then computed, dividing E by the NRCAN precipitation. Since streamflow time series differ from one river basin to another, E/P is calculated only during the available streamflow periods.

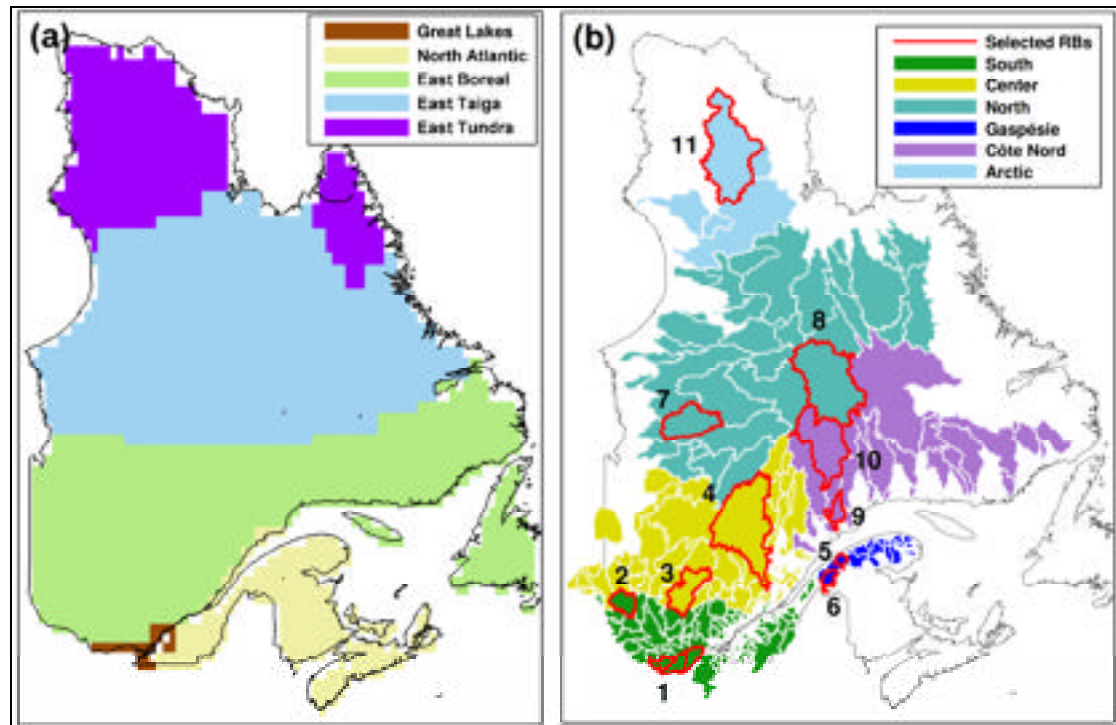


Figure 2.2 (a) Climatic regions of Bukovsky for CFSR, and (b) Hydrologic regions and selected river basins (with red contours) from (cQ)²

2.4 Results

This section reviews the different results obtained during the three analysis steps in this study.

2.4.1 Long-term mean of the hydrological cycle components over the province of Quebec

Figure 2.3 shows the long-term mean of precipitation, evaporation and runoff for the four reanalyses over the province of Quebec. Globally, the spatial distribution of precipitation from ERA-Interim and MERRA are quite similar, whereas precipitation from CFSR generally reaches the highest values among the four datasets (up to 5 mm day⁻¹ over the south of the province). Moreover, the underestimation of precipitation from NARR highlighted by Bukovsky et Karoly (2007) at the US-Canadian border is clearly obvious in Figure 2.3.

Regarding evaporation, MERRA shows higher values in the center and in the south relatively to the other three reanalyses (2.5 to 3 mm day^{-1}), while NARR and ERA-Interim show the lowest values in the north of the province (0 to 1 mm day^{-1}). Over Manicouagan Lake (located at the blue spot of the NARR runoff in the center of the province), CFSR and NARR show high runoff values, especially NARR, which reveals values above 3 mm day^{-1} , while the three other reanalysis runoffs range from 0.5 to 2.5 mm day^{-1} . Furthermore, ERA-Interim shows the highest runoff values in the center and the south, while runoff values from MERRA and NARR drop to almost zero in the west and south, respectively.

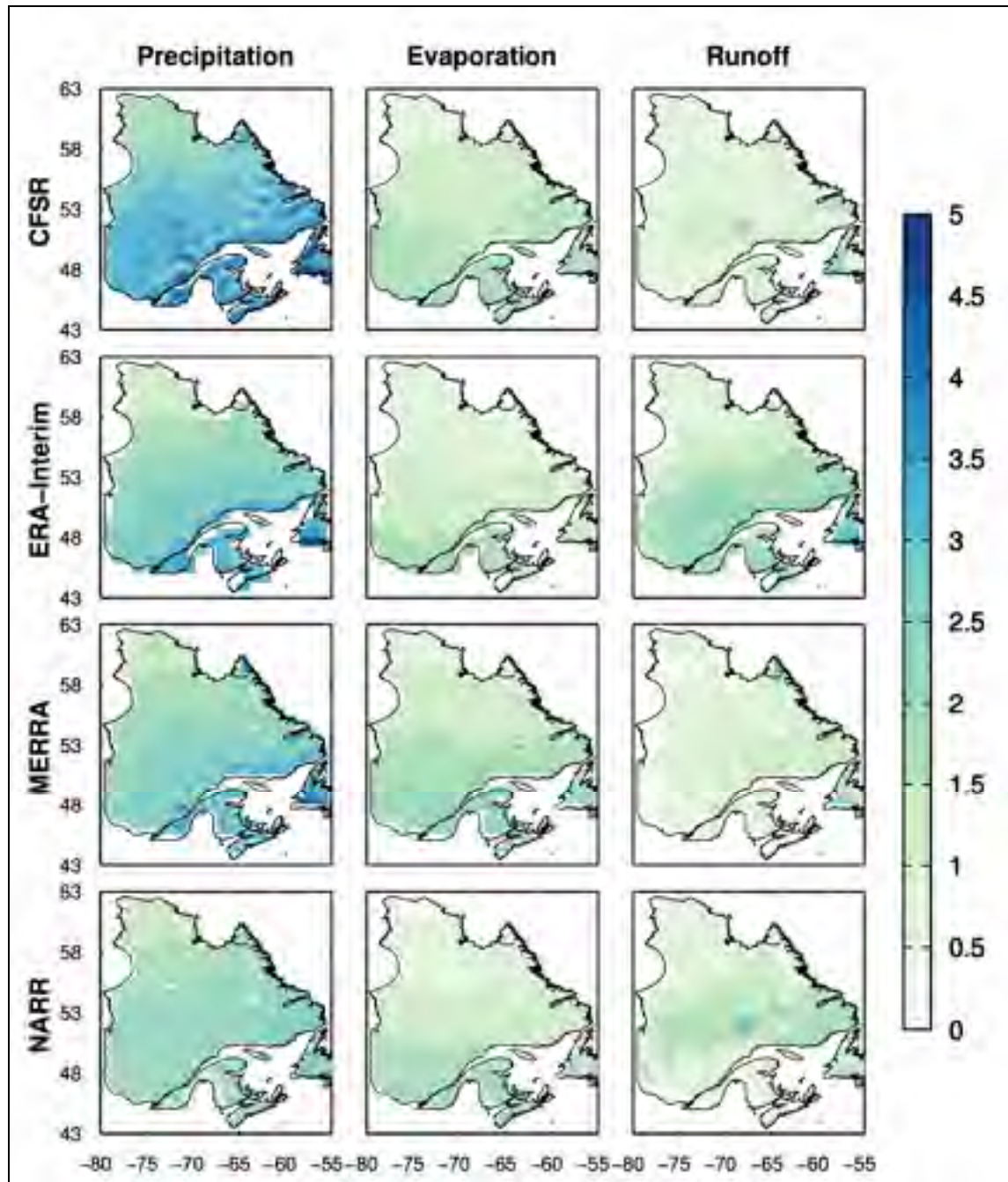


Figure 2.3 Long-term mean precipitation, evaporation and runoff (mm day⁻¹) of CFSR, ERA-Interim, MERRA and NARR over the province of Quebec

Figure 2.4 shows the closure of the water balance using the long-term mean B values computed for CFSR, ERA-Interim, MERRA and NARR over Quebec. As expected, the water balance of MERRA is closed, with values of B almost equal to 0. Conversely, the

water balances of CFSR, NARR and ERA-Interim are not closed, with values of B different from 0. Over the entire province, CFSR presents positive values of B from 30 to 50%. These noticeable results are mainly related to the high values of CFSR precipitation illustrated in Figure 2.3. It is worth recalling that CFSR uses observed precipitation instead of that simulated to force its surface scheme, which introduces some imbalance between evaporation, runoff and the model-generated precipitation. Over central and southern Quebec, ERA-Interim shows B values between -30% and -10%, and a relatively closed water balance with B close to 0 over the rest of the province. NARR presents positive B values of about 20% in the far north latitudes, negative B values in the center, and reaches its lowest negative values in the south. The underestimations of NARR precipitation illustrated in Figure 2.3 at the US-Canadian border and over the Manicouagan Lake lead to significantly underestimated B values (from -70% to -90%).

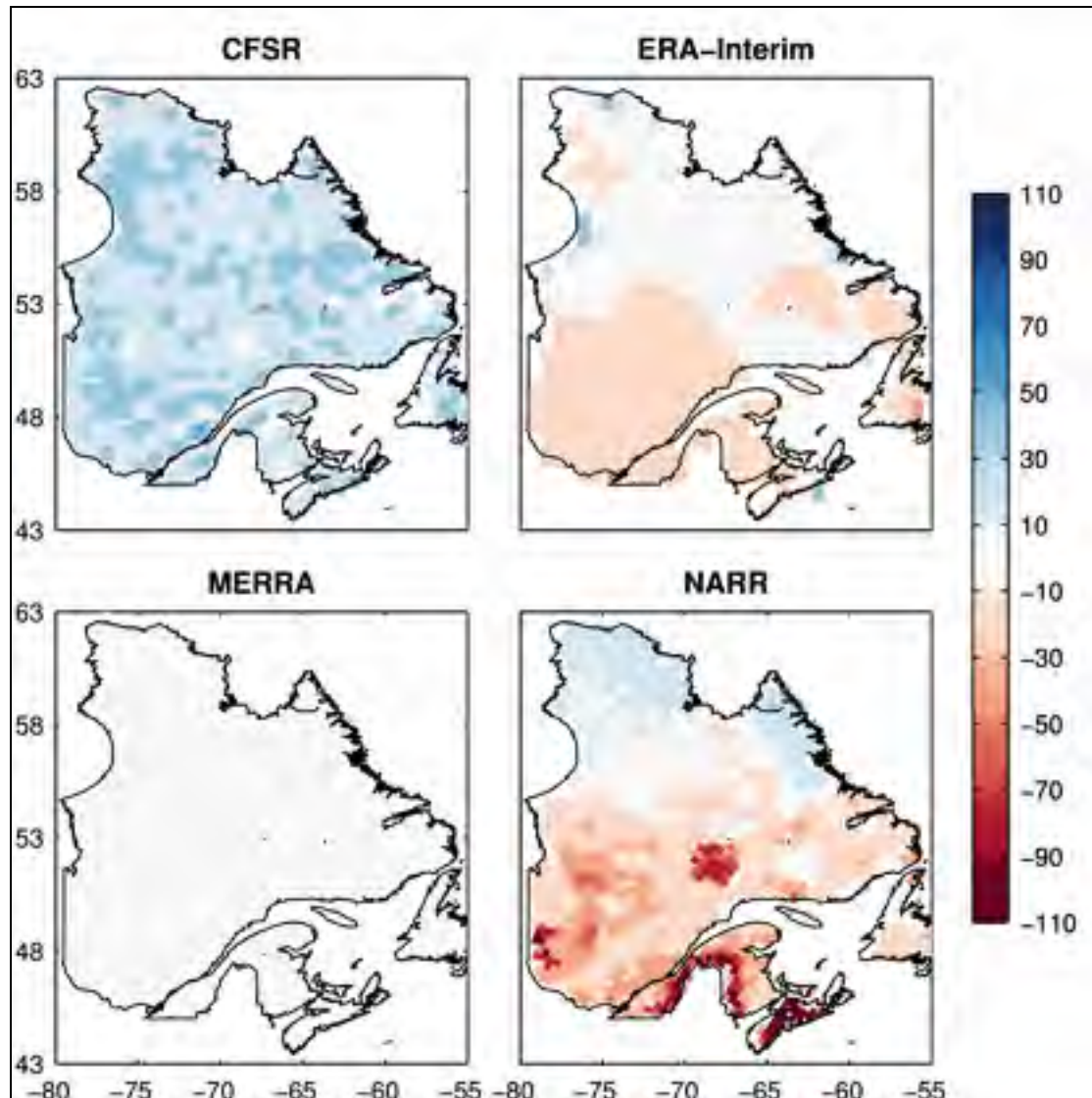


Figure 2.4 Long-term mean relative water balances using B values (% precip) from CFSR, ERA-Interim, MERRA and NARR over the province of Quebec

2.4.2 Hydrological cycle components over the climatic regions

Figure 2.5 shows relative water balance B , runoff, evaporation and precipitation of the four reanalyses over the five climatic regions of Bukovsky, with the addition of the NRCAN dataset for the case of precipitation. Mean annual cycles are shown for runoff, evaporation and precipitation, whereas mean annual values are shown for the relative water balance B (one value per year), from 1979 to 2008.

As already seen in Figure 2.3, precipitation from CFSR is generally higher than that from the other four datasets over the five climatic regions, except in the summer months over the GL and NA regions (Figure 2.5). Precipitation from NARR is lower than for MERRA, ERA-Interim and CFSR, especially over the EB, GL and NA regions ($0,5$ to 1 mm day^{-1} less than the other reanalyses). Nevertheless, although precipitation values may differ greatly depending on the dataset, the temporal distributions are quite similar over the EB, ETA and ETU regions, which is in agreement with the precipitation from NRCAN. Over the GL and NA regions, precipitations from each dataset show discrepancies with each other, mainly in the summer (JJA months).

Regarding evaporation, MERRA summer peak values are systematically higher than those from the other reanalyses, particularly over the EB, ETA and ETU regions, reaching $4,3$, $3,5$ and 3 mm day^{-1} respectively. Over the NA region, the evaporation values from MERRA are also very close to those from NARR, with summer peaks being $1,4$ times higher than those from ERA-Interim and CFSR. Globally, the evaporation is questionable, as summer maximum values may double or triple from one reanalysis to another, depending on the climatic region. The general North-South gradients seen in Figure 2.3 for precipitation and evaporation are also reproduced in the mean annual cycles, as precipitation and evaporation generally tend to decrease from southern to northern regions, with precipitation from CFSR being an exception, as already pointed out, and precipitation over the GL region being lower than that over the NA region. Such gradients agree well with the known gradients in precipitation and evaporation across the province, also illustrated by Fisheries and Environment Canada (1978).

Runoff values from ERA-Interim are higher than those from MERRA, NARR and CFSR all year long for NA and GL regions, and during fall and/or winter seasons for EB, ETA and ETU regions (1 to 2 mm day^{-1} more than the other reanalyses). For instance, over the GL region, the ERA-Interim peak values during the spring are $2,5$ times higher than those from the other three datasets. On the other hand, MERRA, NARR and CFSR agree well over this region. Likewise, MERRA, NARR and CFSR show similar runoff values over the NA

region, whereas those for ERA-Interim are substantially higher. Moreover, MERRA presents higher maxima over ETA and ETU (4,3 and 3,6 mm day⁻¹ respectively) compared with ERA-Interim, NARR and CSFR values. Generally, runoff values from the different reanalyses show more disagreement over the EB, ETA and ETU regions, as the low and high flow seasons show a large range of values.

The relative water balance values from the four reanalysis datasets cover a range from -50% to +50% of precipitation, except for the NARR dataset in the NA region, which will be discussed further. The high precipitation values and average values of evaporation and runoff from CFSR induce positive values of relative water balance B (about 25%), higher than for MERRA, ERA-Interim and NARR over the six climatic regions. In some years, the underestimation of precipitation from NARR leads to negative values of B over the five regions, especially over NA in the early 2000s, where B values are significantly low, below -150%. This may have been induced by the high evaporation values during summer, combined with the low precipitation of NARR in this region, which includes the US-Canadian border where precipitation from NARR is strongly underestimated (section 2.4.1). Despite the highest evaporation maxima from MERRA, this reanalysis shows a closed water balance with B values around 0% over the five climatic regions, as expected (section 2.3.1). Regarding the ETA and ETU regions, ERA-Interim reveals a relatively closed water balance with B values close to 0, compared with NARR, and particularly CFSR. However, the higher runoff values from ERA-Interim over the EB, GL and NA regions tend to produce a negative relative water balance (about -15%) for most years. Globally, the reanalyses do not agree in terms of relative water balance, while some of their hydrological cycle components present some similarities. On the other hand, although MERRA was designed to close its water balance, it also reveals some overestimation of different hydrological cycle components (with respect to the other reanalyses, and with respect to NRCAN observed precipitation mostly in ETA and ETU regions).

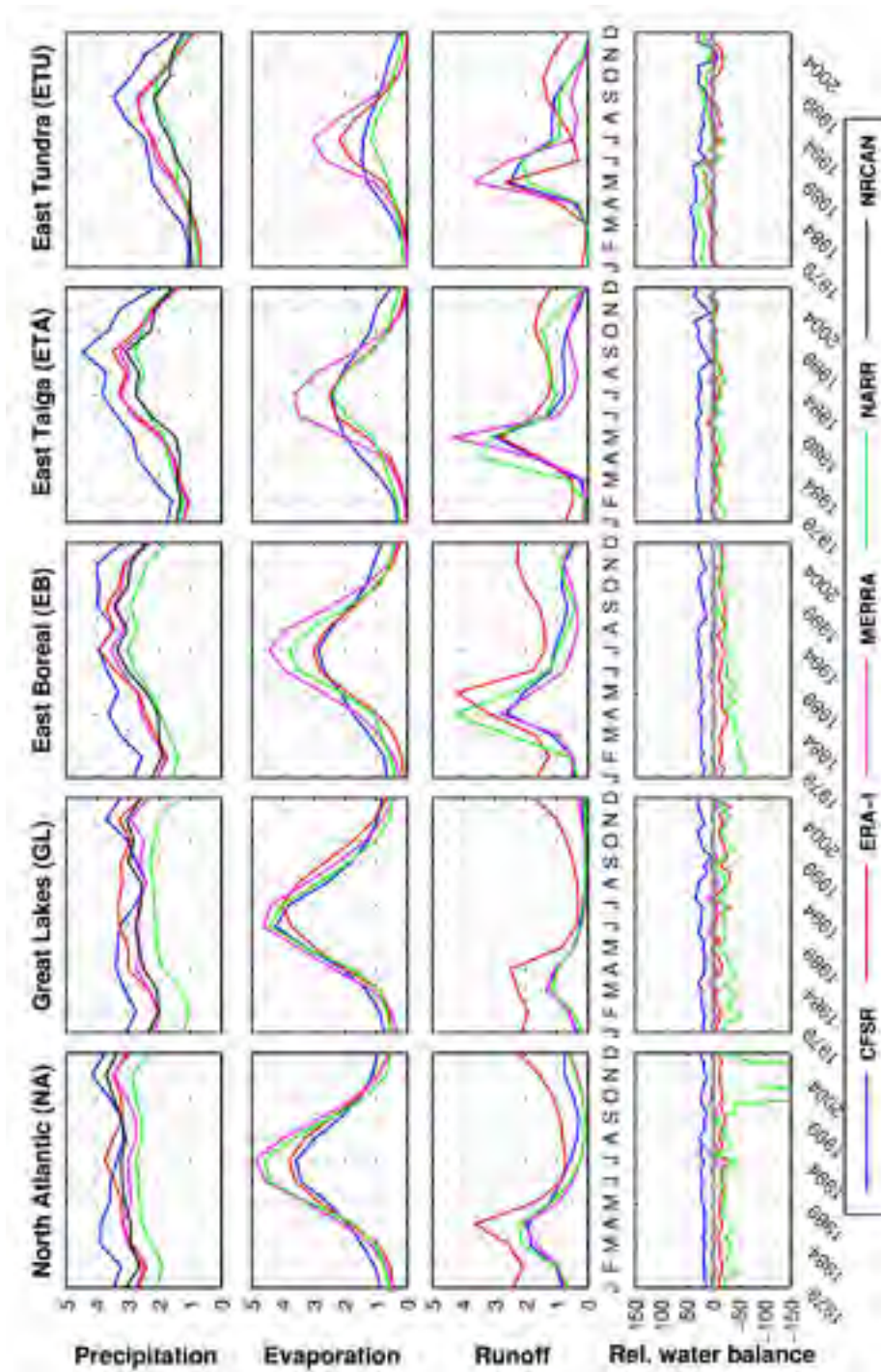


Figure 2.5 Mean annual cycle of precipitation, evaporation and runoff (in mm day^{-1}), and annual mean relative water balance (in % of precip), from 1979 to 2008, of MERRA, ERA-Interim, NARR and CFSR, with the addition of NRCAN for the case of precipitation for the five climatic regions from Bukovsky (2011)

2.4.3 Hydrological cycle components within river basins

Mean annual precipitation from the four reanalyses and NRCAN datasets were computed from 1979 to 2008 (Figure 2.6) as spatial averages over the 11 selected river basins (RB) (section 2.3.3.3 and Figure 2.2 (b)). As already seen in the previous sections, precipitation from CFSR is generally higher than that of the other datasets over the 11 river basins, including NRCAN (Figure 2.6). Among the four reanalyses, MERRA and ERA-Interim seem to be the less biased with respect to the NRCAN observations. On the other hand, NARR presents significant underestimations, as compared to ERA-Interim and MERRA, especially over RB 1 and RB 6 (1,5 to 2 mm day⁻¹ less). Furthermore, a sudden change in precipitation from NARR around 2003 is clearly noticeable over RBs 1 to 7 (which are respectively located in the GL climatic region for RB 1, in the EB region for RB 2, 3 and 4, and in the NA region for RBs 5 and 6).

Looking at Figure 2.7, the spatial distribution of the mean precipitation from NARR is quite different between the 1979-2002 and 2003-2008 periods. Indeed, the bias towards low precipitation values over the US-Canadian border between 1979 and 2002 vanished between 2003 and 2008. In southern Quebec, there is significantly more precipitation after 2003 than before. This could likely be related to the update of the NARR assimilation system in April 2003, and to the change in the number and nature of assimilated observations (Mesinger et al., 2004).

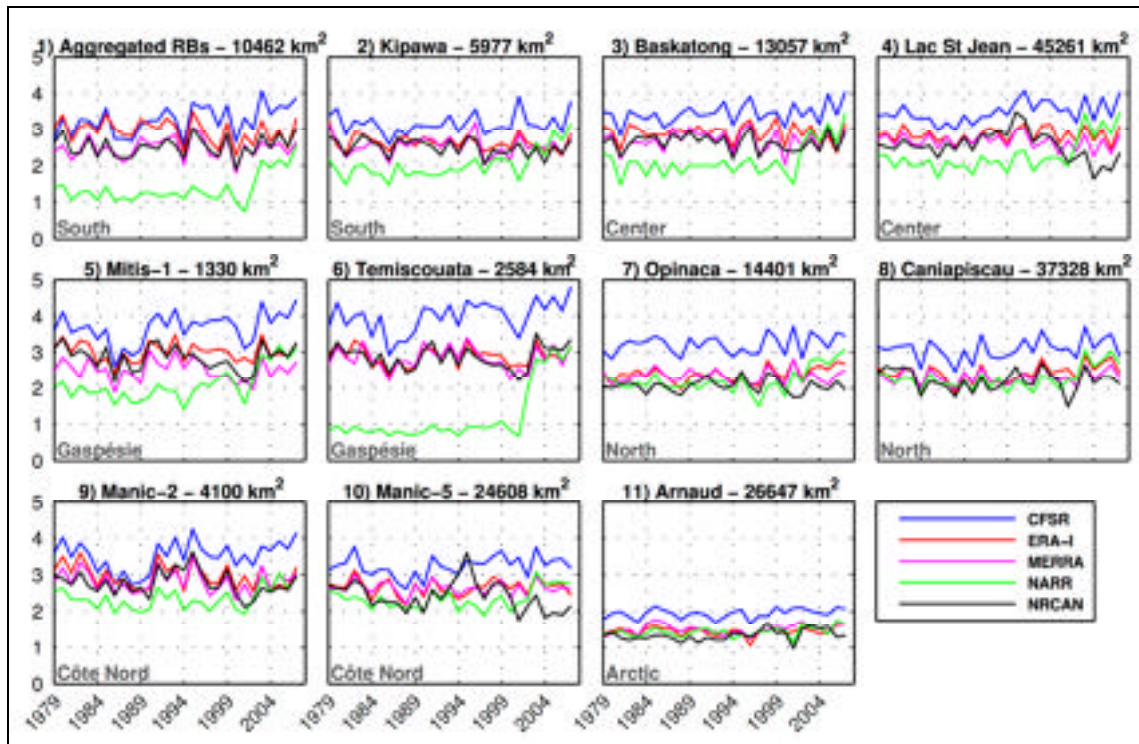


Figure 2.6 Mean annual precipitation (mm day⁻¹) from CFSR, ERA-Interim, MERRA, NARR and NRCAN, from 1979 to 2008 over the 11 selected river basins. For each river basin, the total area is indicated in square kilometers, and the hydrologic region is indicated on the bottom left corner

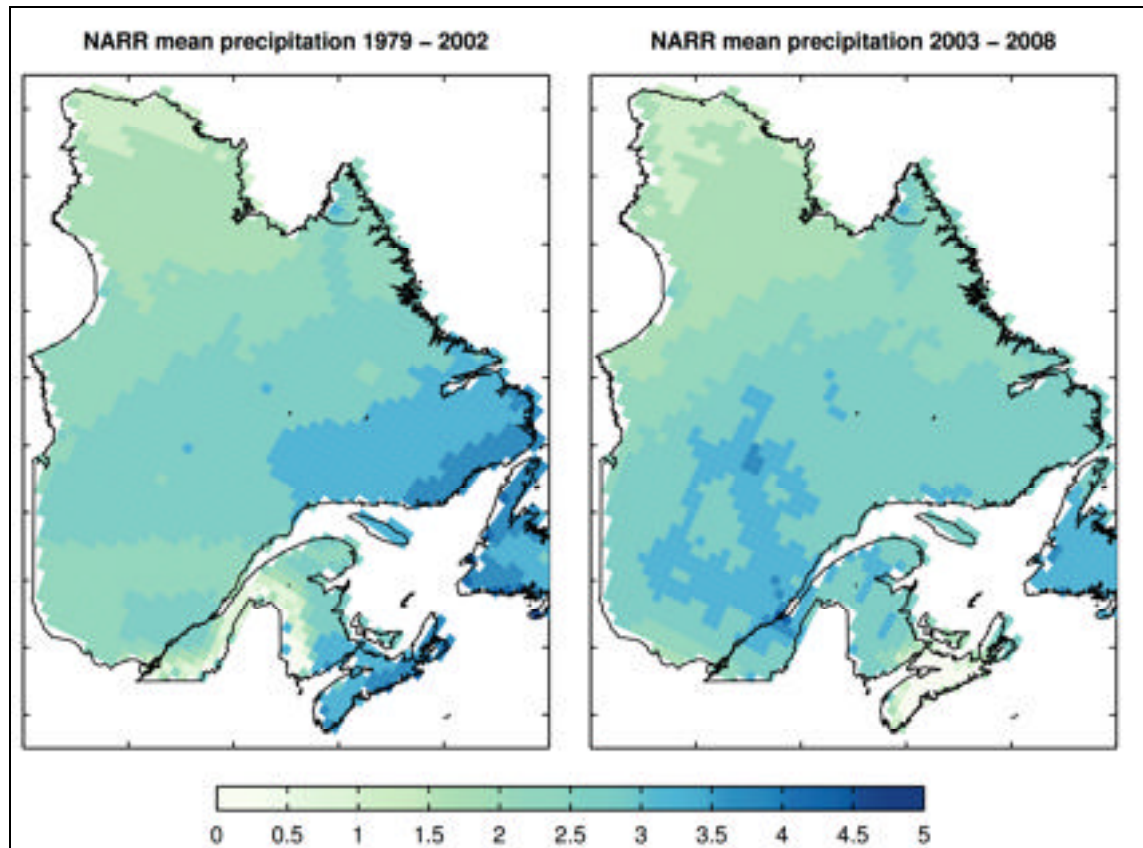


Figure 2.7 Mean precipitation of NARR (mm day^{-1}) for the periods 1979-2002 (left) and 2003-2008 (right), over the province of Quebec

Tableau 2.2 shows the correlation between daily time series of the spatial averaged precipitation from each of the four reanalyses and NRCAN over the selected river basins (section 2.3.3.3). Despite the global overestimation of precipitation from CFSR, the correlations over all river basins are similar to those from MERRA and ERA-Interim. On the other hand, NARR presents low or inconsistent correlations, compared with MERRA, ERA-Interim and CFSR from -0,33 to 0,62. The negative value of NARR over RB 10 (-0,13) is not surprising when considering the low correlations of the other three datasets. However, the negative correlation value from NARR over RB 4 (-0,33) is probably related to the sudden shift in precipitation happening around 2003, discussed above (Figure 2.7). Nevertheless, these results must also be taken with caution since the quality of the precipitation from NRCAN is questionable in the northern part of the province, as the gridded data have been interpolated from very few weather stations (section 2.3.2). For further analysis, it could be

relevant to perform a direct comparison between reanalysis grid points and measurements from the closest weather stations in order to avoid biases induced by the generation of gridded observation datasets. In any case, a correlation between daily time series around 0,8 or more should be acceptable to indicate that reanalyses represent the climate over the river basins of interest quite well, especially in southern Quebec, and therefore, may be useful for hydrological modelling purpose.

Tableau 2.2 Correlation between daily time series of spatially averaged precipitation of CFSR, ERA-Interim, MERRA and NARR versus NRCAN over the 11 selected river basins

	HYDROLOGIC REGION	CFSR	ERA-Interim	MERRA	NARR
(1) Aggregated RBs	South	0.78	0.78	0.81	0.62
(2) Kipawa	South	0.58	0.72	0.73	0.13
(3) Baskatong	Center	0.69	0.57	0.81	0.42
(4) Lac St-Jean	Center	0.34	0.45	0.62	-0.33
(5) Mitis-1	Gaspésie	0.75	0.80	0.79	0.41
(6) Temiscouata	Gaspésie	0.74	0.81	0.74	0.50
(7) Opinaca	North	0.36	0.26	0.16	0.09
(8) Caniapiscau	North	0.65	0.56	0.59	0.41
(9) Manic-2	Côte Nord	0.60	0.67	0.71	0.33
(10) Manic-5	Côte Nord	0.31	0.15	0.23	-0.13
(11) Arnaud	Arctic	0.14	0.13	0.04	0.37
Average		0.54	0.53	0.57	0.26

Figure 2.8 shows mean annual runoff from CFSR, ERA-Interim, MERRA, NARR and the observational estimates over the 11 river basins. The streamflow records for the 11 selected river basins do not completely cover the 1979-2008 period, so the runoff calculation periods differ among the river basins. Despite the high values of runoff from ERA-Interim over GL, NA, and EB climatic regions (Figure 2.3), this reanalysis show runoff values similar to the runoff estimates at the watershed scale, except over the RB 3,4 and 9. On the other hand, CFSR and MERRA reveal comparable amount of runoff water (from 0,5 to 1 mm .day⁻¹) over all RB, and systematically lower than ERA-Interim and the runoff estimates (about 1 to

1,5 mm day⁻¹ less). Over RB 10, CFSR and mostly NARR present high pic runoff values (3 and 5,5 mm day⁻¹ respectively) on the period 1979 - 1984. This river basin encompasses the Manicouagan Lake, where high local values of runoff were also highlighted on Figure 2.3 for the case of NARR and CFSR. As this result is common to these two reanalyses, one can assume that their common surface model Noah may be the cause of those high runoff values.

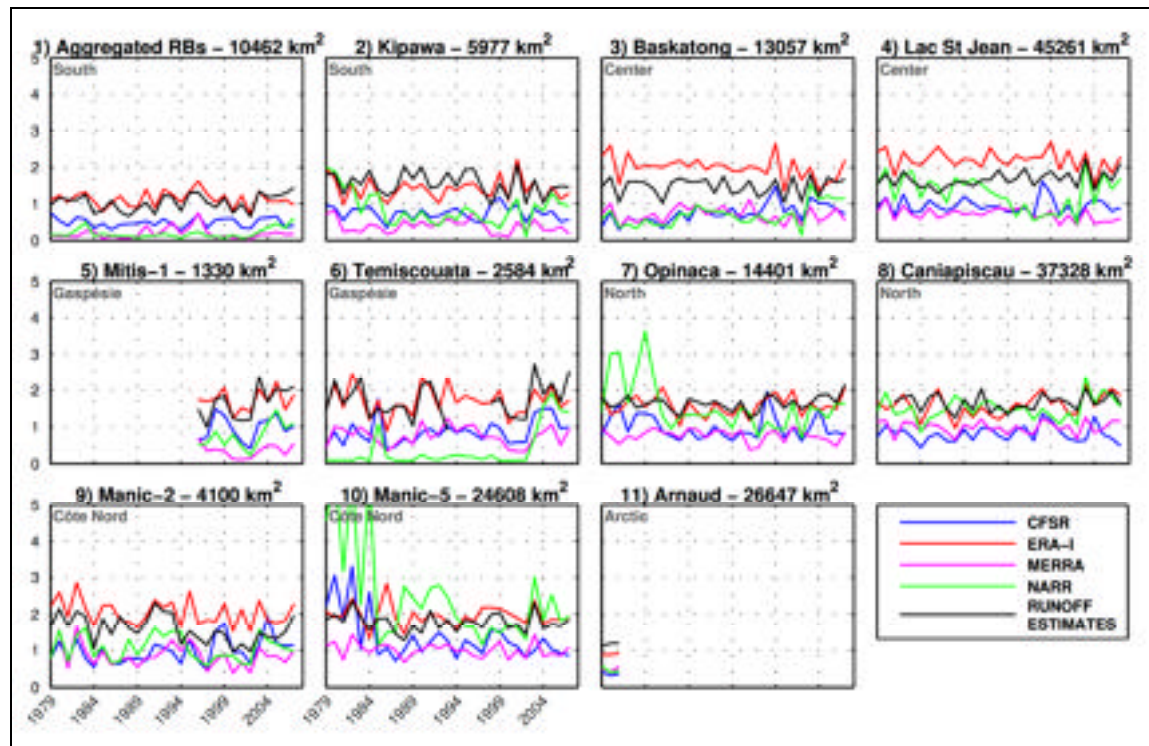


Figure 2.8 Mean annual runoff (mm day⁻¹) from CFSR, ERA-Interim, MERRA, NARR and NRCAN, from 1979 to 2008 over the 11 selected river basins. Annual time series are limited to the available streamflow observations

Figure 2.9 shows mean annual E/P ratios from CFSR, ERA-Interim, MERRA, NARR and the observational estimates over the 11 river basins. As for the runoff results, the E/P calculation periods differ among the river basins. Considering the quality of the precipitation from NRCAN and of some (cQ)² streamflow time series, one can expect the observational estimates to sometimes be questionable (section 2.3.2). Therefore, the observational estimates should not always be considered as the absolute truth, but rather, as another dataset for the purpose of analysis.

Over southern and central Quebec, evaporation is expected to be about a half of precipitation, which means an E/P ratio around 0,5 (Fisheries and Environnement Canada, 1978; Wang et al. 2014). Looking at Figure 2.9, the observational estimates ratio approaches 0,5 over RBs 1 and 3, and moderately underestimates 0,5 over RB 2. Looking more closely at RBs 4 and 10, the observational estimates ratio falls in the early 2000s. This fall is also noticeable in Figure 2.6, where, starting in 2004, the precipitation from NRCAN shows a slight downward trend, related to the quality issue of this precipitation dataset (discussed in section 2.3.2) over Lac St-Jean and west of Côte Nord region, where RBs 4 and 10 are respectively located.

For all river basins, MERRA and NARR show the highest E/P ratio values, especially NARR, which even exceeds 1 over the South, Center and Gaspésie river basins, which is not physically consistent. The high ratio values from NARR over RB 6 (and to a smaller extent, RBs 1 and 2) are directly related to the underestimation of precipitation close to the US-Canadian border, which extends through the river basins of Gaspésie (Figure 2.2 (b) and Figure 2.7). Ratio values from MERRA tend to (and even sometimes exceed) 1 for the river basins in the South, Gaspésie and Center regions. This means that almost all the precipitation evaporates, which implies that there is little water left for runoff in these regions for MERRA. This particularity is also illustrated in Figure 2.5 where MERRA shows the lowest runoff and the highest evaporation over the southern climatic regions in summer. Rienecker et al. (2011) highlighted deficiencies in MERRA that lead to the immediate evaporation of much of the rainfall, and consequently limit the surface runoff. Regarding ERA-Interim and CFSR, the E/P ratios are quite consistent with the expected value of 0,5 over the South and Center regions, except for some higher values in RB 1 between 1994 and 2008. Globally, the E/P ratios of the reanalyses show a more steady evolution during the 30 years over the North and Côte Nord river basins versus in the South, Gaspésie and Center river basins. However, the observational estimates ratios are most often lower than for the four reanalyses in the northern regions. In fact, the general limitations (discussed in section 2.3.2) of the observed streamflow and the precipitation time series reduce the credibility of the evaporation observational estimates in the northern part of the province. Over RB 11 in the Arctic region,

only three years of streamflow data were available to compute the E/P ratio. Nevertheless, the results show that the general behaviours of the reanalyses and the observational estimates appear to be similar to those over the other northern river basins.

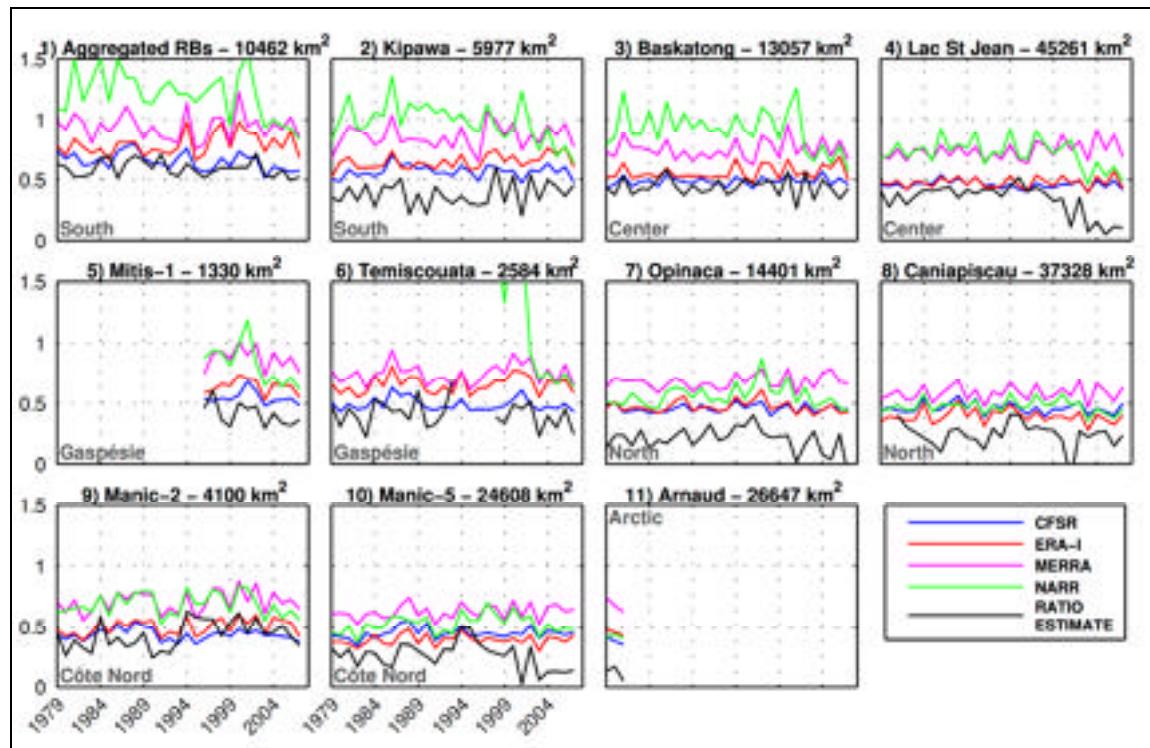


Figure 2.9 Annual E/P ratio (evaporation divided by precipitation) from CFSR, ERA-Interim, MERRA, NARR and observational estimates from 1979 to 2008 over the 11 selected river basins. Annual time series are limited to the available streamflow observations that were used to estimate E from observations

2.5 Concluding remarks

The scope of this study was to determine the reliability of the CFSR, MERRA, ERA-interim and NARR reanalyses in representing the terrestrial branch of the water cycle over the province of Quebec and to explore their potential in providing meteorological variables where few or no observations are available. The long-term mean of the water balance and the hydrological cycle components were investigated for the four reanalyses. A first look at the hydrological cycle components of the reanalyses showed that the water balance is not always

maintained. Globally, the MERRA water balance is closed, as compared to the other three reanalyses, whereas NARR and ERA-Interim revealed negative water balances in central and southern Quebec, and CFSR revealed a positive water balance. The negative water balance of NARR is strongly influenced by inconsistencies in its components: for instance, underestimations of precipitation over the US-Canadian border and a large high bias of evaporation over Lac Manicouagan. Furthermore, the high precipitation from CFSR compared with those of MERRA, ERA-Interim and NARR, contributes to the positive values of the water balance over the entire province.

Precipitation, evaporation and runoff were also assessed on a multi-year monthly basis, over the Bukovsky climatic regions (Bukovsky, 2011). For the four reanalyses, mean annual cycles of the hydrological cycle components showed a similar temporal evolution. However, the amounts of water differed, especially the maxima of evaporation and runoff, and the precipitations from NARR and CFSR. A closer look at precipitation from the four reanalyses at the river basin scale confirmed that values from ERA-interim and MERRA agreed fairly well with one another and with those of NRCAN. On the other hand, CFSR and NARR almost systematically represented the highest and the lowest precipitation values, respectively, and consequently, do not appear to be particularly reliable over the province of Quebec, especially in the south.

Nevertheless, the temporal correlation of the daily precipitation from the reanalyses versus the NRCAN over the river basins showed that precipitation from CFSR, ERA-Interim and MERRA vary quite synchronously despite water amount disparities between them. Precipitation from NARR, which showed an overall poor daily correlation, should be taken with care, depending on the region of interest. Despite its highest runoff values among the reanalyses over southern climatic regions, ERA-Interim showed runoff values similar to runoff estimates over almost all studied river basins. Moreover, The E/P ratios revealed that ERA-Interim and CFSR succeeded relatively well in reproducing the distribution of precipitated water into evaporation and runoff, whereas NARR showed physically inconsistent results over southern Quebec. Regarding MERRA, although this dataset was

designed to close the water balance, its surface model does not properly distribute an adequate amount of precipitated water into evaporation and runoff, since most of the precipitated water seems to evaporate before running off. Therefore, each of the four reanalyses has its own strengths and weaknesses. NARR, which has been shown to be quite reliable in the continental United States (Bukovsky et Karoly, 2007; Mesinger et al., 2004; Sheffield, Livneh et Wood, 2012), appeared to be the least consistent dataset among the four reanalyses over eastern Canada. MERRA and ERA-Interim provide probably the most reliable precipitation over the province of Quebec, whereas evaporation and runoff remain questionable, except for the runoff of ERA-Interim, which shows the most consistent values, compared with the observations.

No matter the case, for impact studies, care should be taken in using any of the reanalysis terrestrial hydrological cycle components, for instance, in using evaporation and runoff as reference data in water resource management or as input or calibration data when performing hydrological modelling. In fact, it could be relevant to compare reanalysis datasets with weather station measurements (in the case of precipitation) to validate the ability of reanalyses to represent the weather, especially with the small number of weather stations available in northern regions. A next step to this study will involve feeding hydrological models with reanalysis temperature and precipitation, and evaluating and comparing their potential for model calibration and validation. Hydrological variables from reanalyses could also be used as reference data in evaluating intermediate steps of simulation within the hydrological model (evaporation, aquifer height, snowmelt), or introduced inside hydrological models in order to reduce the number of calibration parameters. Nevertheless, even though further investigations are necessary, these reanalyses have good potential to provide meteorological and hydrological information in remote areas of Canada. Moreover, some variables of these reanalyses succeed in revealing consistent values in regions where observational datasets are reliable.

2.6 Acknowledgements

We thank the National Centers for Environmental Prediction, the National Center for Atmospheric Research, the National Aeronautics and Space Administration and the European Centre for Medium-Range Weather Forecasts for developing and providing the reanalysis datasets. CFSR and NARR were obtained through the CISL Research Data Archive website at rda.ucar.edu/pub/cfsr.html and rda.ucar.edu/datasets/ds608.0, respectively. MERRA was downloaded from gmao.gsfc.nasa.gov/merra and ERA-Interim was downloaded from apps.ecmwf.int. We also thank Catherine Guay from the Institut de Recherche d'Hydro-Québec for providing the (cQ)² database and useful additional information about these data. We also would like to thank Blaise Gauvin St-Denis from Ouranos for providing valuable help in downloading and formatting reanalysis datasets, and finally, the organizations that funded this project, the Conseil de Recherches en Sciences Naturelles et en Génie du Canada, Hydro-Québec, Rio Tinto Alcan and Ontario Power Generation.

CHAPITRE 3

ÉTUDE DE LA PERFORMANCE DES RÉANALYSES EN MODÉLISATION HYDROLOGIQUE

Suite aux travaux d'étude du cycle hydrologique des réanalyses sur le Québec (chapitre 2), ce chapitre décrit les travaux d'évaluation de la performance des réanalyses en modélisation hydrologique. Les résultats obtenus apporteront de l'information quant à l'applicabilité des réanalyses pour la simulation des débits en rivière, notamment dans le nord du Québec, où les données d'observation sont rares. La première section de ce chapitre présente la méthodologie adoptée pour le calage et la validation du modèle hydrologique HSAMI. La seconde section présente les résultats obtenus à l'aide du modèle HSAMI, et la troisième section introduit les discussions et conclusions des travaux de modélisation hydrologique.

3.1 Méthodologie

Cette section expose les processus de préparation des données de réanalyses et d'observations, ainsi que les méthodes de calage et validation du modèle hydrologique HSAMI.

3.1.1 Préparation des données

Cette section présente la méthodologie adoptée pour préparer les données utilisées lors des processus de calage et validation du modèle hydrologique HSAMI. Les principales caractéristiques des réanalyses et des données d'observations ne sont pas décrites ici. Pour obtenir de l'information sur les données utilisées, se référer aux sections 2.3.1 et 2.3.2 du chapitre 2.

3.1.1.1 Bassins versants de (cQ)²

L'étude de la performance en modélisation hydrologique des réanalyses CFSR, MERRA, ERA-Interim et NARR, ainsi que du jeu de données d'observation sur grille NRCAN a été réalisée avec le modèle global conceptuel HSAMI, développé par Hydro-Québec. La base de données (cQ)² de séries temporelles de débits observés a été choisie comme référence pour le calcul de l'indice de performance en calage et en validation. La base de données (cQ)² contient des séries de débits observés sur 306 bassins versants au Québec, dont certains s'étendent en Ontario ou aux États-Unis (Arsenault et Brissette, 2014). La plage de superficie des 306 bassins versants s'étend de 10 à 70 000 km². La création de cette base de données a été entreprise par Hydro-Québec (HQ), le Centre d'Expertise Hydrique du Québec (CEHQ) et Rio Tinto Alcan (RTA), et contient donc des informations sur des bassins versants gérés par les trois gestionnaires (Figure 3.1).

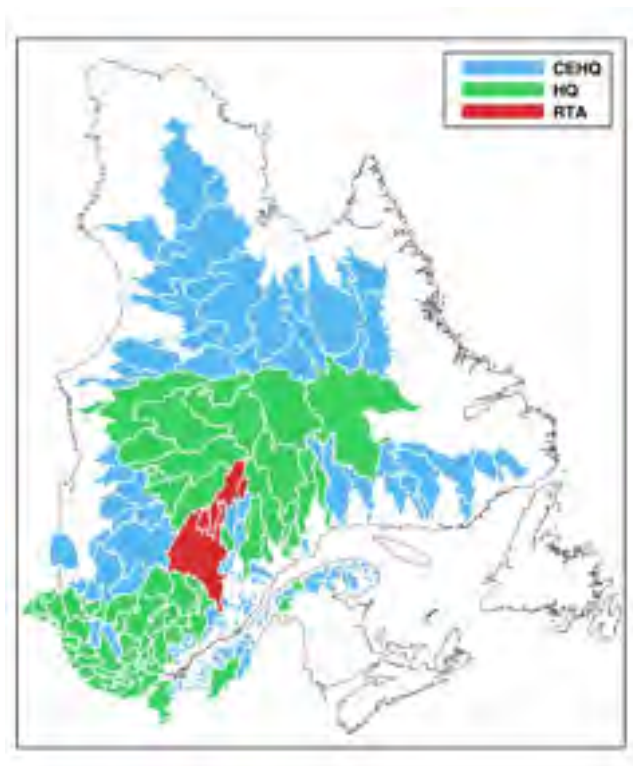


Figure 3.1 Localisation des 306 bassins versants de (cQ)² et leurs gestionnaires

Certaines séries de débits ont fait l'objet de traitement: les débits naturels ont été reconstitués sur les bassins régulés, certaines séries d'étiages hivernaux ont été simplement retirées à cause des erreurs de mesure dues aux surpressions induites par les embâcles, d'autres ont été complétées par des apports simulés par HSAMI pour combler les trous (communication personnelle Catherine Guay, Institut de recherche d'Hydro-Québec, Février 2015).

La qualité des séries de débits reste variable, et compte tenu des modifications apportées, il apparaît donc nécessaire de trier et filtrer les bassins versants en fonction de leur série de débits, avant d'effectuer le calage du modèle HSAMI. L'objectif est de conserver, parmi les 306 bassins versants de $(cQ)^2$, les bassins dont la série de débits observés s'étend sur la période 1979-2008 (recouvrement temporel commun avec les réanalyses) et présente une qualité acceptable pour calculer l'indice de performance hydrologique. De plus, les bassins versants seront étudiés selon les régions hydrologiques du Québec, précédemment définies dans le chapitre 2. Il est donc nécessaire de conserver suffisamment de bassins pour représenter chacune des six régions hydrologiques. Ainsi, une série de filtres, comportant différents critères, a été appliquée par l'intermédiaire d'un algorithme. La Figure 3.2 présente le processus de filtrage effectué sur le total de 306 bassins versants.

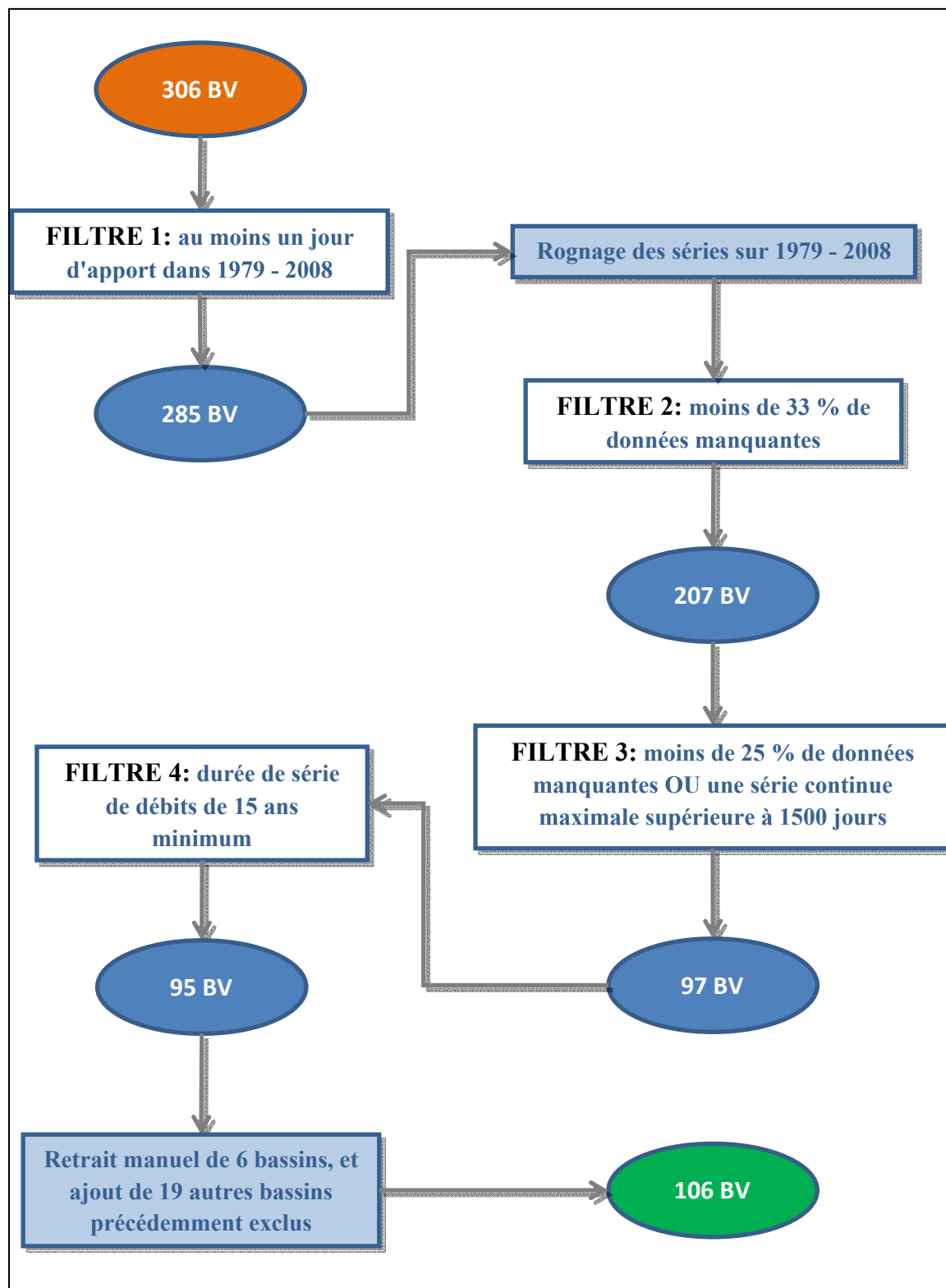


Figure 3.2 Schéma de l'algorithme de filtrage des bassins versants de $(cQ)^2$. Les critères des filtres indiqués dans les boîtes concernent les bassins versants qui sont conservés

Le premier filtre conserve uniquement les bassins versants dont la série de débits possède au moins un jour en commun avec la période 1979-2008. Cette période correspond à la période de recouvrement de 30 ans choisie pour les réanalyses, et c'est aussi sur cette période que l'étude du cycle hydrologique des réanalyses a été menée. Les séries de débits des bassins conservés sont ensuite rognées sur cette période.

Le second filtre conserve uniquement les bassins versants dont la série de débits contient moins de 33% de données manquantes. Dans $(cQ)^2$, beaucoup de bassins versants ne présentent aucune donnée lors des étiages. De plus, les données manquantes sont éparpillées sur toute la durée de la série de débits pour la grande majorité des bassins. Le second filtre permet donc de conserver les bassins présentant des données lors des étiages.

Le troisième filtre conserve les bassins versants dont la série de débits contient moins de 25% de données manquantes OU une série continue maximale supérieure à 1500 jours. Ainsi, le troisième filtre permet d'affiner la sélection du second filtre, mais aussi de conserver les bassins qui présenteraient une série continue suffisamment longue pour être exploitée en modélisation hydrologique, malgré le grand nombre de données manquantes présentes dans cette série.

Le quatrième filtre conserve les bassins versants dont la série de débits présente une durée totale de 15 ans minimum. Il est préférable d'avoir un nombre suffisant d'année sur lesquelles l'algorithme d'optimisation calculera l'indice de performance. En effet, plus la durée sur laquelle le modèle sera calé est longue, plus le paramétrage du modèle sera robuste. En d'autres termes, avec une telle combinaison de paramètres, le modèle devrait être capable de reproduire un panel plus large de comportements hydrologiques (différentes intensités et durées de crues et d'étiages, par exemple).

Suite au quatrième filtre, un nombre de 6 bassins versants a été retiré manuellement, car leur série de débits a été jugée pas assez fiable pour leur utilisation dans cette étude. Par ailleurs, dans le but d'étudier suffisamment de bassins versants dans chacune des régions

hydrologiques du Québec, il a été décidé de rajouter 19 bassins précédemment exclus lors des différents filtrages. Par exemple, le bassin versant Aux Feuilles, dont la série de débits observés ne dépasse pas 10 ans, a été récupéré pour représenter la région Arctique. Finalement, le nombre total de bassins versants conservés s'élève à 106. La Figure 3.3 présente la localisation de ces 106 bassins versants selon leur région hydrologique. Les contours en rouge correspondent aux douze bassins versants dont les cycles moyens annuels des débits simulés et observés sont analysés plus en détails dans la section 3.2.3. Chaque bassin versant est rattaché à une région hydrologique. Ces régions découpent le territoire du Québec selon les régimes hydrologiques majeurs de la province. D'autre part, les principales rivières d'une même région hydrologique s'écoulent dans une même direction, vers un cours d'eau ou plan d'eau commun. Par exemple, les rivières des bassins versants de la région Côte Nord s'écoulent du Nord au Sud vers le fleuve Saint-Laurent, alors que les rivières des bassins versants de la région Nord s'écoule d'Est en Ouest vers la baie d'Hudson. Ainsi, la région Sud compte 36 bassins, la région Centre compte 30 bassins, la région Gaspésie compte 11 bassins, la région Nord compte 11 bassins, la région Côte Nord compte 17 bassins et la région Arctique compte 1 bassin (Figure 3.3).

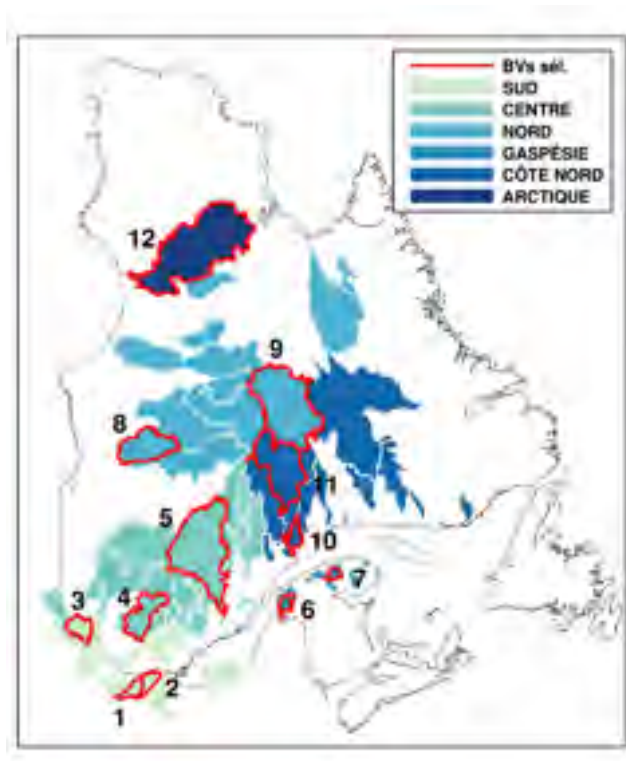


Figure 3.3 Localisation des 106 bassins versants conservés après filtrage, selon les régions hydrologiques du Québec. Les 12 bassins versants dont les cycles annuels moyens des débits simulés et observés sont analysés dans la section 3.2.3, sont indiqués par un contour rouge

3.1.1.2 Réanalyses et données d'observations sur grille

Les précipitations et températures minimales et maximales des réanalyses CFSR, MERRA, ERA-Interim, NARR et du jeu de données d'observation sur grille NRCAN ont été agrégées spatialement en un point et au pas de temps journalier sur chacun des 106 bassins versants sélectionnés sur la période 1979 à 2008. Dans le cas de NARR, uniquement la température moyenne est disponible. Les températures minimales et maximales de NARR ont donc été estimées de la manière suivante: les températures minimales et maximales du point de grille NRCAN le plus proche de chaque point de grille NARR sont récupérées. La moitié de la différence entre températures maximales et minimales de NRCAN est ajoutée ou retranchée à la température moyenne de NARR à chaque point de grille pour créer respectivement les températures maximales et minimales de NARR. Les Figure-A II-1 et Figure-A II-2 en Annexe II présentent les moyennes sur la période 1979 - 2008 des températures minimales et

maximales de NRCAN, CFSR, ERA-Interim et MERRA sur les territoires de l'Est du Canada. À première vue, les gradients des températures minimales et maximales de NRCAN et ERA-Interim sont relativement similaires, alors que les températures minimales de MERRA et les températures maximales de CFSR présentent un biais froid vis-à-vis de celles de NRCAN dans la moitié Nord de la province. Par ailleurs, MERRA présentent un L'agrégation des données de précipitation et température par bassin versant a été effectuée selon la méthode suivante: lorsqu'au moins quatre points de la grille de données se situent à l'intérieur du contour du bassin, une moyenne simple est utilisée. Sinon, la méthode des Polygones de Thiessen décrite dans la section 2.3.3.3 est appliquée (Rhynsburger, 1973). De plus, l'algorithme d'agrégation des données a été optimisé pour qu'aucun point de grille des réanalyses au-dessus de la mer ne soit pris en compte. Ceci a été réalisé en appliquant le masque terre-mer de chacune des réanalyses.

3.1.2 Calage et validation du modèle HSAMI

HSAMI est un modèle global conceptuel à 23 paramètres (Fortin, 2000) développé par Hydro-Québec en 1983. Depuis, il a fait l'objet de nombreuses améliorations, et est toujours utilisé par les gestionnaires pour la prévision hydrologique opérationnelle sur différents bassins versants se trouvant sur le territoire du Québec. Ce modèle a été utilisé à plusieurs reprises par le passé dans le cadre d'études hydroclimatiques (Arsenault et al., 2013a; Chen, Brissette et Leconte, 2011; Minville, Brissette et Leconte, 2009; Poulin et al., 2011; Riboust, 2014).

Basé sur des réservoirs liés par des relations linéaires, HSAMI modélise les processus hydrologiques majeurs ayant lieu sur un bassin versant: accumulation et fonte de neige, interception, infiltration, écoulement vertical, écoulements horizontaux, évapotranspiration. Le gel de la zone non-saturée et la hauteur de l'aquifère sont aussi simulés par ce modèle. Les réservoirs du modèle sont: la neige au sol, le réservoir de surface, le gel, les zones non-saturée et saturée (sol), et une réserve intermédiaire. Le modèle inclut la possibilité, le cas échéant, d'inclure la présence d'un lac ou d'un réservoir sur le territoire d'un bassin étudié.

Pour plus de détails sur l'architecture du modèle et son fonctionnement, se référer à Fortin (2000).

Les données fournies en entrée du modèle HSAMI sont la précipitation, ainsi que les températures minimale et maximale, sur une base journalière. HSAMI peut prendre en compte des données auxiliaires comme l'ensoleillement ou le pourcentage de la superficie du bassin occupée par le réservoir (ou lac), mais ces données ne sont pas utilisées dans cette étude. Habituellement, les précipitations neigeuses et pluvieuses (sous forme d'équivalent en eau) doivent être séparées avant d'être fournies au modèle. Un module de séparation de la pluie et de la neige a donc été ajouté en amont du modèle. Ce module contient deux paramètres ajustables, assimilables aux températures minimales et maximales en-deçà/au-delà desquelles la précipitation sera totalement solide/liquide. Ces deux paramètres s'ajoutent aux 23 paramètres du modèle qui seront optimisés lors du calage.

L'algorithme d'optimisation choisi pour calibrer le modèle HSAMI est le Covariance Matrix Adaptation Evolution Strategy CMAES (Hansen et Ostermeier, 1996; 2001), suivant les indications de Arsenault et al. (2013b). Le coefficient de Nash-Sutcliffe (NSE) a été sélectionné comme indice pour évaluer la performance du modèle en calage et en validation. Cet indice présente la particularité de dépendre en grande partie des débits de crue, plutôt que des étiages (Moriassi et al., 2007; Pushpalatha et al., 2012). Selon la durée des 106 séries de débits observés, le NSE n'est pas nécessairement calculé sur la même période pour chaque bassin versant, la période maximale étant 1979-2008. Le calage du modèle est réalisé sur les années paires, et la validation sur les années impaires, pour les cinq jeux de données: CFSR, MERRA, ERA-Interim, NARR et NRCAN. Un total de 30 calages a été effectué avec chaque jeu de données, sur chaque bassin versant. Ceci permet de prendre en compte l'équifinalité des paramètres en calage, ainsi qu'une éventuelle variabilité dans l'optimisation effectuée par l'algorithme CMAES. Le jeu de paramètres correspondant au meilleur NSE en calage a été retenu pour effectuer la validation, pour chaque bassin et chaque jeu de données.

3.2 Résultats

Cette section présente tout d'abord les résultats en calage du modèle HSAMI, avec les cinq jeux de données d'entrées NRCAN, NARR, ERA-Interim, CFSR et MERRA. Par la suite, les résultats en validation sont analysés, puis une étude de corrélation entre la superficie des bassins versants et la performance des réanalyses en modélisation est examinée.

3.2.1 Calage

La Figure 3.4 représente les distributions des NSE en calage sur les 106 bassins versants sélectionnés, pour NRCAN, NARR, ERA-Interim, CFSR et MERRA. La figure inclut deux boîtes à moustaches pour chaque jeu de données: la première correspond à la distribution des NSE pour les 30 calages de chaque bassin, la seconde correspond à la distribution des meilleurs NSE en calage pour chaque bassin (donc une valeur par bassin). Par exemple, dans la région Sud et pour chaque jeu de données, la première boîte à moustaches inclut un total de 1080 points (30 calages x 36 bassins), et la seconde inclut 36 points (36 bassins). Sur chaque boîte à moustaches, le point central représente la médiane et les arêtes de la boîte représentent les 25^e (P_{25}) et 75^e (P_{75}) percentiles. Les horsains sont définis par défaut à l'extérieur de l'intervalle $P_{25} - 1.5*(P_{75} - P_{25})$ et $P_{75} + 1.5*(P_{75} - P_{25})$.

Globalement, les calages avec les cinq jeux de données sont satisfaisants, révélant des valeurs de Nash-Sutcliffe généralement supérieures à 0,6. Dans les régions Sud et Gaspésie, les données NRCAN performant mieux que les quatre réanalyses, alors que dans les régions plus au nord (Centre, Nord, Côte Nord et Arctique), les réanalyses ERA-Interim, CFSR et MERRA performant mieux que NRCAN, malgré leurs résolutions grossières. Rappelons ici que les données sur grille NRCAN sont interpolées sur un nombre de plus en plus faible de stations vers le nord, voire presque nul dans le grand nord. La qualité des données NRCAN dans le nord est donc probablement impactée par le manque de données d'observation.

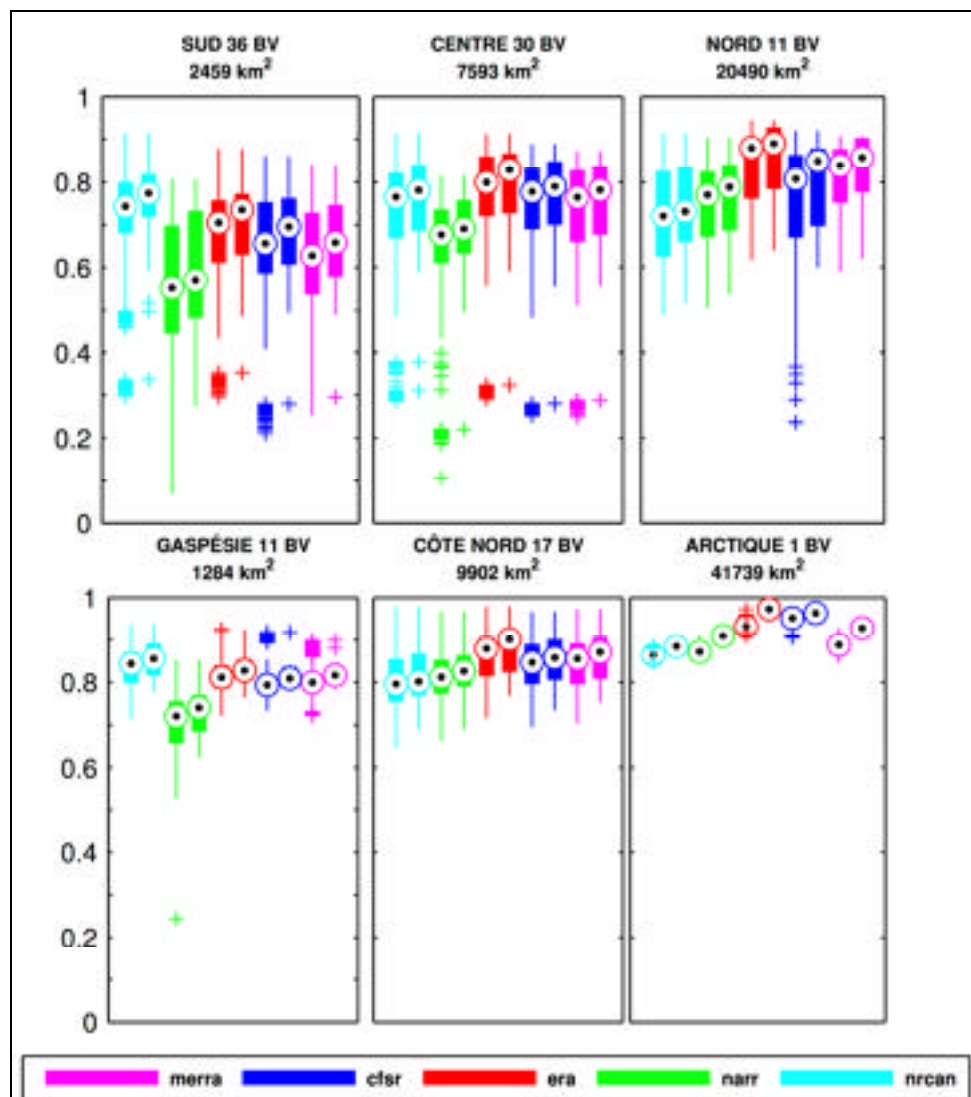


Figure 3.4 Distribution par région hydrologique des coefficients de Nash-Sutcliffe (en ordonnée) pour les jeux de données MERRA, CFSR, ERA-Interim, NARR et NRCAN. Il y a deux boîtes à moustaches pour chaque jeu de données: la première correspond à la distribution des NSE pour les 30 calages de chaque bassin, la seconde correspond à la distribution des meilleurs NSE pour chaque bassin (une valeur par bassin). Le titre de chaque sous-figure indique le nombre de bassins versants dans chaque région hydrologique, ainsi que la superficie moyenne des bassins versants

D'autre part, des groupements de points sont observables dans la distribution des NSE des régions Sud et Centre. En réalité, ces horsains correspondent aux mêmes bassins versants pour les cinq jeux de données, dans chacune des deux régions. Ainsi, par exemple, pour les bassins versants Lac St-François et Gouin (régions Sud et Centre respectivement), les valeurs

de NSE varient entre 0,2 et 0,3, et pour les bassins versants Mountain Chute et Madawaska-Arnprior (région Sud), entre 0,4 et 0,5, et cela, pour les cinq jeux de données. La Figure 3.5 représente les séries de débits observés de quatre bassins versants, sur lesquels les performances des cinq jeux de données sont systématiquement mauvaises. Ces séries révèlent de fortes irrégularités sur les 30 ans: une brusque variation dans le régime annuel est observable aux environs de 1992-1993 pour les bassins versants Lac St-François et Gouin, et aux environs de 1996-1997 pour les bassins versants Mountain Chute et Madawaska-Arnprior. Malgré la procédure de sélection qui a été appliquée sur la totalité des bassins versants de $(cQ)^2$, il y a donc des séries de débits peu exploitables qui sont passées à travers les filtres de sélection. Ceci appuie donc d'autant plus l'importance de vérifier la qualité des données d'observations avant d'évaluer la performance des réanalyses en modélisation. Par exemple, un test d'homogénéité pourrait être ajouté pour compléter la série de filtre déjà appliqués, de manière à retirer les séries de débits présentant des caractéristiques similaires à celles de ces quatre bassins versants.

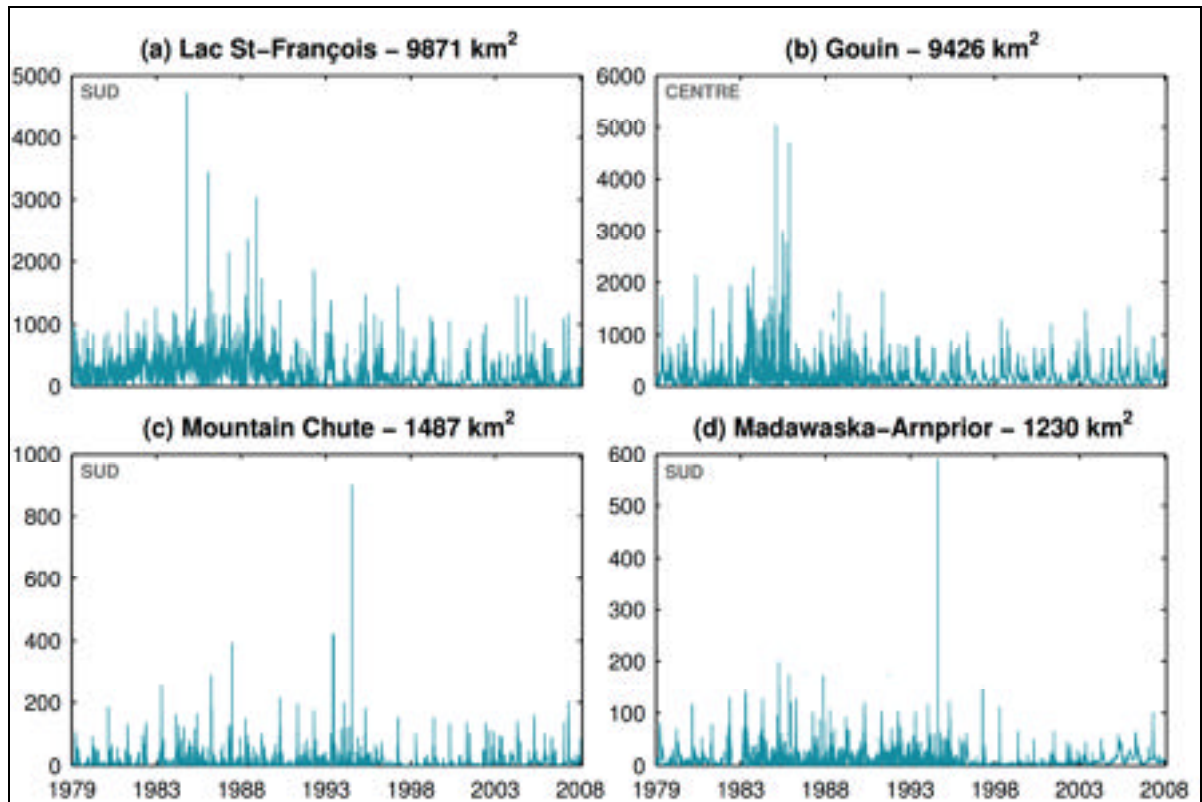


Figure 3.5 Séries des débits observés journaliers sur les bassins versants (a) Lac St-François, (b) Gouin, (c) Mountain Chute et (d) Madawaska-Arnprior (en $\text{m}^3 \text{s}^{-1}$). La région hydrologique est indiquée en haut à gauche et la superficie de chaque bassin versant dans le titre des sous-figures

3.2.2 Validation

La Figure 3.6 représente la distribution des NSE en validation sur les 106 bassins versants sélectionnés pour NRCAN, NARR, ERA-Interim, CFSR et MERRA. Chaque distribution est constituée d'un point par bassin, soit par exemple, 36 points pour la région Sud et 1 point pour la région Arctique. La première correspond aux NSE ayant été produits par les jeux de paramètres générant le meilleur NSE parmi les 30 calages. La seconde correspond aux meilleurs NSE parmi les 30 validations. L'intervalle défini pour les valeurs centrales et les horsains est le même que dans la Figure 3.4.

La Figure 3.6 indique en premier lieu que les paramètres ayant produit le meilleur NSE en calage ne produisent pas nécessairement le meilleur NSE en validation (première distribution versus seconde distribution, pour chaque jeu de données). Toutefois, le gain en performance reste relativement minime. L'analyse des résultats qui suit s'adresse, en premier lieu, aux premières distributions de NSE pour chaque jeu de données, mais de façon plus générale, s'appliquent également aux secondes distributions de la Figure 3.6.

Les résultats de validation sont légèrement décalés à la baisse par rapport aux résultats en calage, à l'exception des résultats sur la région Arctique. Dans cette région, hormis la meilleure performance de MERRA en validation, les performances des quatre autres jeux de données ont baissé, notamment celles de NARR et de CFSR qui ont fortement chuté. Ces résultats mettent en lumière la relation entre la durée de la période au cours de laquelle le modèle est calibré, et la robustesse des jeux de paramètres. En effet, la série de débits observés du bassin versant Aux Feuilles (Arctique) s'étalent sur huit ans (quatre années pour le calage et quatre années pour la validation) sans compter les données journalières manquantes de la série, alors que les autres séries de débits ont une durée de 30 ans pour la plupart. En conséquence, les jeux de paramètres plus robustes (calés sur 15 ans), révèlent généralement une diminution du NSE plus faible en validation que celles des jeux de paramètres calés sur quatre ans.

Dans les régions Centre, Nord, Gaspésie et Côte Nord, les valeurs médianes des NSE varient entre 0,6 et 0,8, pour CFSR, MERRA, ERA-Interim et NRCAN. Les performances de ces quatre jeux de données sont sensiblement identiques dans ces régions, avec une légère supériorité non significative (voir la discussion qui suit à propos des différences statistiquement significatives entre les distributions des NSE) d'ERA-Interim dans les régions Centre, Nord et Côte Nord et de NRCAN en Gaspésie et dans le Sud.

Dans la région Sud, les performances sont moins élevées, avec des valeurs de NSE, à l'intérieur des écarts interquartiles s'étalant entre 0,3 et 0,75. D'autre part, dans les régions Sud, Centre et Gaspésie, NARR présente les distributions de NSE les plus faibles. La sous-

estimation de la précipitation de NARR le long de la frontière américaine, ainsi que la brusque altération du régime annuel de ses précipitations se produisant au cours de l'année 2003 (discutées dans la section 2.4) sont probablement à l'origine de ces résultats. D'autre part, parmi les quatre réanalyses, ERA-Interim, dont la résolution est la plus grossière (un point de grille pour environ 7000 km²) montre des valeurs de NSE égales, voire supérieures aux valeurs de NSE de CFSR et NARR (un point de grille pour environ 1000 km²). La résolution plus fine d'une réanalyse ne contribue donc pas nécessairement à une performance supérieure à celle d'une réanalyse dont la résolution est plus grossière. Toutefois, bien que la différence de résolution soit assez significative entre les réanalyses étudiées, leurs résolutions demeurent grossières vis-à-vis de l'échelle spatiale des plus petits bassins versants parmi les 106 sélectionnés (de l'ordre de quelques dizaines ou quelques centaines de kilomètres carrés).

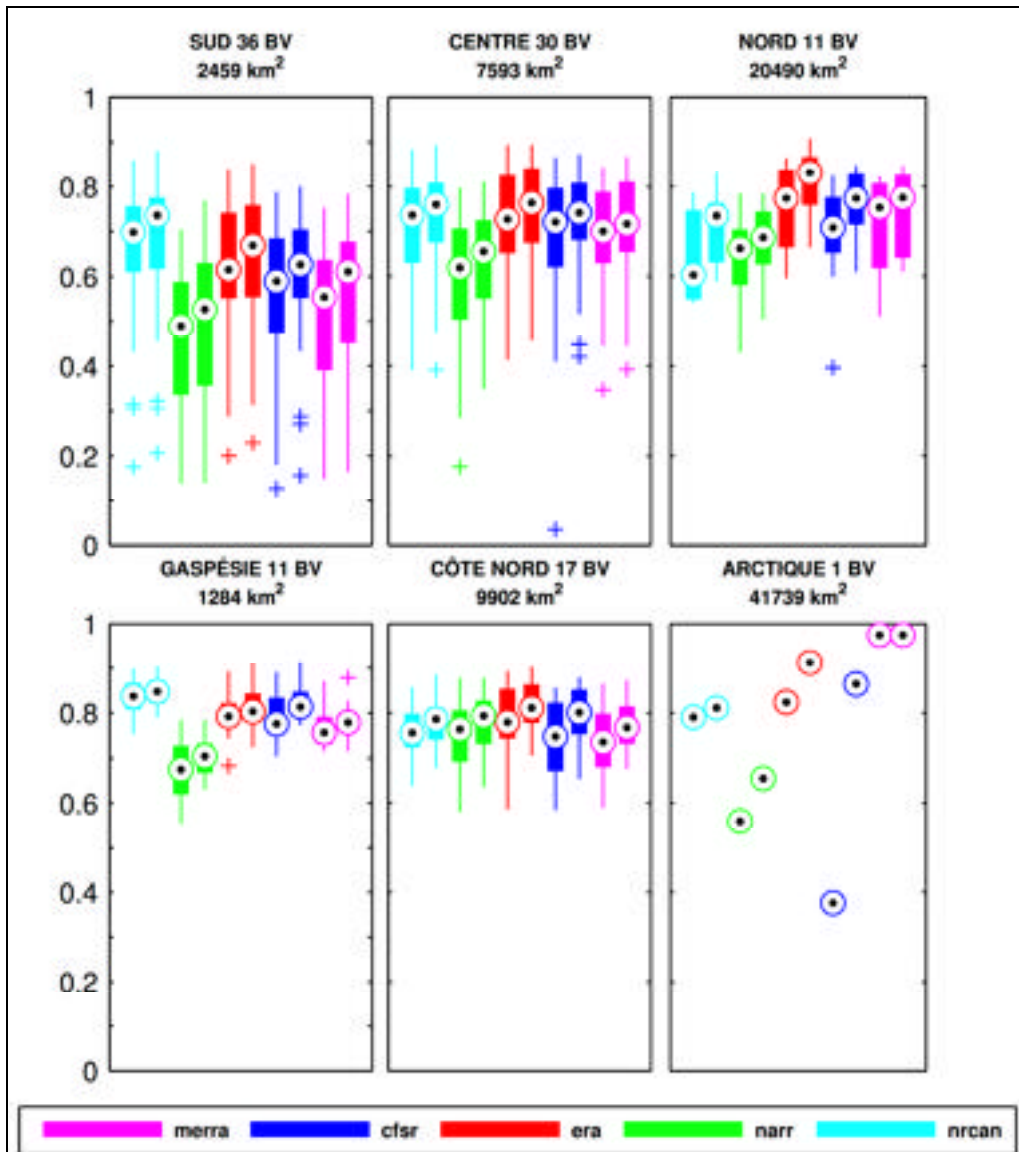


Figure 3.6 Distribution par région hydrologique des coefficients de Nash-Sutcliffe (en ordonnée) en validation pour les jeux de données MERRA, CFSR, ERA-Interim, NARR et NRCAN. Il y a deux boîtes à moustaches pour chaque jeu de données: la première correspond aux NSE ayant été produits par les jeux de paramètres générant le meilleur NSE parmi les 30 calages. La seconde correspond aux meilleurs NSE parmi les 30 validations

Par ailleurs, un test de comparaison statistique multiple a été appliqué entre les distributions des NSE des cinq jeux de données (seconde boîte à moustaches) sur les régions Sud, Centre, Nord, Gaspésie et Côte Nord. Ce test, appelé test de Kruskal-Wallis (Kruskal et Wallis, 1952), repose sur l'hypothèse nulle (H_0) selon laquelle les échantillons sont issus de la même

population. L'autre hypothèse (H1) stipule qu'on ne peut conclure que tous les échantillons proviennent de la même population. Il s'agit d'un test non-paramétrique, basé sur les rangs moyens des échantillons.

Ce test est appliqué sur les cinq jeux de données par l'intermédiaire de la fonction MATLAB *multcompare*, qui permet de visualiser la différence statistique d'un échantillon avec un autre grâce à l'étude des rangs moyens des points des échantillons. Cette fonction produit donc une figure représentant les rangs moyens des échantillons analysés, ainsi que leur intervalle de confiance. Deux échantillons sont significativement différents si leurs intervalles de confiance sont disjoints. Au contraire, si leurs intervalles de confiance se chevauchent, les échantillons ne sont pas significativement différents. Ainsi, les Figure-A I-1, Figure-A I-2, Figure-A I-3, Figure-A I-4 et Figure-A I-5 en Annexe I présentent pour les régions Sud, Centre, Nord, Gaspésie, et Côte Nord respectivement, les valeurs des rangs moyens (cercles) et son intervalle de confiance (trait) des distributions de NSE en calage (correspondant aux secondes boîtes à moustaches sur la Figure 3.4) de chaque jeu de données. Le rang moyen est calculé à partir des valeurs des rangs de chaque NSE parmi la totalité des NSE (donc parmi les cinq jeux de données).

Ces résultats montrent dans un premier temps que la distribution des NSE de MERRA n'est statistiquement différente de celles des trois autres réanalyses sur aucune région. ERA-Interim révèlent des distributions de NSE statistiquement différentes de NARR sur les régions, Sud, Centre et en Gaspésie. De même, NARR et CFSR révèlent aussi des distributions de NSE statistiquement différentes en Gaspésie. Finalement, NRCAN montrent des distributions de NSE statistiquement différentes de NARR et de MERRA dans la région Sud et en Gaspésie, et de NARR dans la région Centre. Pour plus de détails sur ces résultats, se référer aux figures en Annexe I.

3.2.3 Analyses complémentaires

La Figure 3.7 représente l'hydrogramme annuel moyen des débits simulés par NRCAN, NARR, ERA-Interim, CFSR et MERRA ainsi que l'hydrogramme annuel moyen de la série de débits observés de 12 bassins versants, notamment certains bassins versants sélectionnés dans l'analyse du cycle hydrologique des réanalyses dans le chapitre 2. Ces 12 bassins versants sont répartis sur les six régions hydrologiques et sont délimités en rouge sur la Figure 3.3. Le Tableau 3.1 présente les coefficients de Nash-Sutcliffe obtenus avec NRCAN, NARR, ERA-Interim, CFSR et MERRA en validation sur les douze bassins versants sélectionnés (rappel: validation réalisée sur les années impaires de 1979 à 2008 avec le jeu de paramètres produisant le meilleur NSE en calage). Les faibles performances de NARR exposées précédemment sur les bassins versants du Sud et de la Gaspésie, sont notables sur la Figure 3.7 pour les bassins 1, 2 et 6, où la réanalyse présente les étiages hivernaux et les pics moyens de crues printanières les plus bas. Par ailleurs, les cinq jeux de données surestiment les étiages estivaux sur les bassins du Sud (bassins 1, 2 et 3). Pour les bassins 8, 10 et 11 (régions Nord et Côte Nord), l'ensemble des réanalyses tend à sous-estimer les pointes de crues annuelles moyennes. Toutefois, les pics de crues semblent relativement bien reproduits sur les bassins de la région Centre (4 et 5), de la Gaspésie (excepté NARR sur les bassins 6 et 7) et sur le bassin 9 (Côte Nord). Les comportements les plus erratiques, entre les différentes réanalyses, sont notés pour le bassin 12 (région Arctique) pour les raisons évoquées précédemment (section 3.2.2).

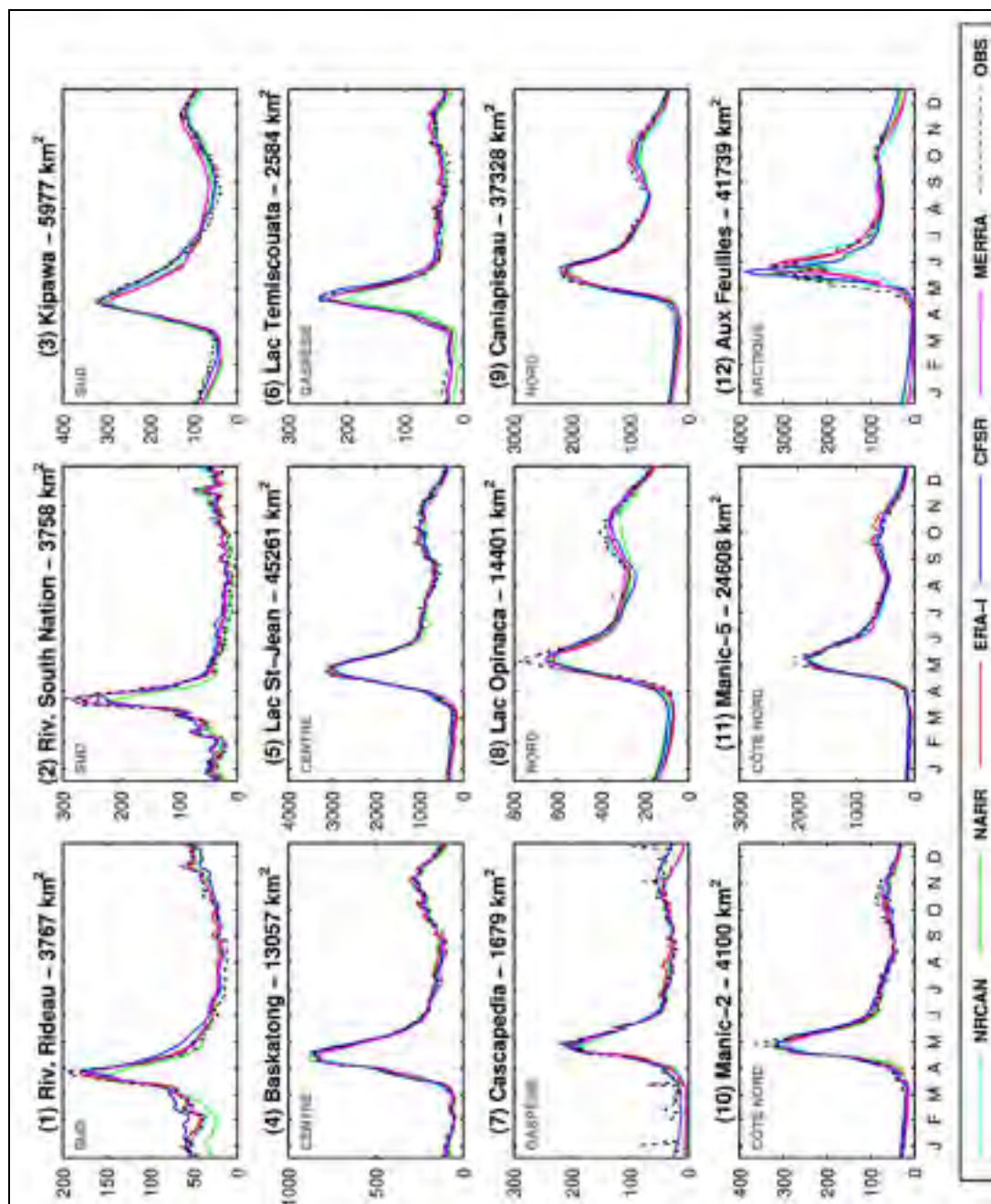


Figure 3.7 Hydrogrammes moyens annuels des débits simulés par NRCAN, NARR, ERA-Interim, CFSR et MERRA et des débits observés sur 12 bassins versants du Québec. Les débits sont indiqués en ordonnées ($\text{m}^3 \text{s}^{-1}$) et les mois en abscisse

Tableau 3.1 Performances (NSE) du modèle HSAMI avec NRCAN, NARR, ERA-Interim, CFSR et MERRA en validation sur 12 bassins versants du Québec. Les moyennes par bassin versant et par jeu de données sont indiquées respectivement dans la dernière colonne et dernière ligne

Bassins versants	NRCAN	NARR	ERA-Interim	CFSR	MERRA	Moyenne
(1) Riv. Rideau (Sud - 3767 km ²)	0.69	0.46	0.69	0.61	0.63	0.62
(2) Riv. South Nation (Sud - 3758 km ²)	0.78	0.45	0.72	0.55	0.55	0.61
(3) Kipawa (Sud - 5977 km ²)	0.80	0.57	0.73	0.74	0.68	0.70
(4) Baskatong (Centre - 13057 km ²)	0.79	0.71	0.84	0.80	0.77	0.78
(5) Lac St-Jean (Centre - 45261 km ²)	0.82	0.75	0.85	0.84	0.82	0.82
(6) Temiscouata (Gaspésie - 2584 km ²)	0.84	0.61	0.80	0.76	0.73	0.75
(7) Cascapedia (Gaspésie - 1679 km ²)	0.83	0.67	0.78	0.83	0.77	0.78
(8) Lac Opinaca (Nord - 14 410 km ²)	0.75	0.71	0.77	0.77	0.75	0.75
(9) Caniapiscau (Nord - 37 328 km ²)	0.79	0.76	0.86	0.78	0.80	0.80
(10) Manic-2 (Côte Nord - 4100 km ²)	0.75	0.76	0.78	0.75	0.67	0.74
(11) Manic-5 (Côte Nord - 24 608 km ²)	0.74	0.75	0.86	0.77	0.87	0.80
(12) Aux Feuilles (Arctique - 41 739 km ²)	0.79	0.56	0.82	0.38	0.97	0.70
Moyenne	0.78	0.65	0.79	0.72	0.75	0.74

La Figure 3.8 représente les coefficients de Nash-Sutcliffe des bassins versants sélectionnés classés selon la superficie des bassins. Les bassins versants dont les séries de débits observés sont inférieures à 9000 jours (au nombre de 7) ont été écartés de cette figure. En effet, l'objectif de cette analyse est de déterminer l'existence d'une éventuelle corrélation entre la performance des réanalyses en modélisation hydrologique, et la superficie des bassins versants modélisés. Or, certaines séries de débits ne comportent qu'un nombre restreint de données, par exemple, le bassin versant Aux Feuilles (8 années), et dont le NSE calculé pour chaque réanalyse pourrait fausser l'analyse. Sur cette figure, les droites de régression linéaire (en rouge sur chaque sous-figure) des quatre réanalyses présentent une pente plus importante que celle de la droite de régression linéaire de NRCAN. Ainsi, le lien entre la valeur du coefficient de Nash-Sutcliffe et la superficie des bassins versants semble un peu plus marqué pour les réanalyses que pour NRCAN. Toutefois, il serait nécessaire de compléter cette analyse de manière à pouvoir exprimer clairement l'existence d'une corrélation entre la performance des réanalyses et la superficie des bassins versants.

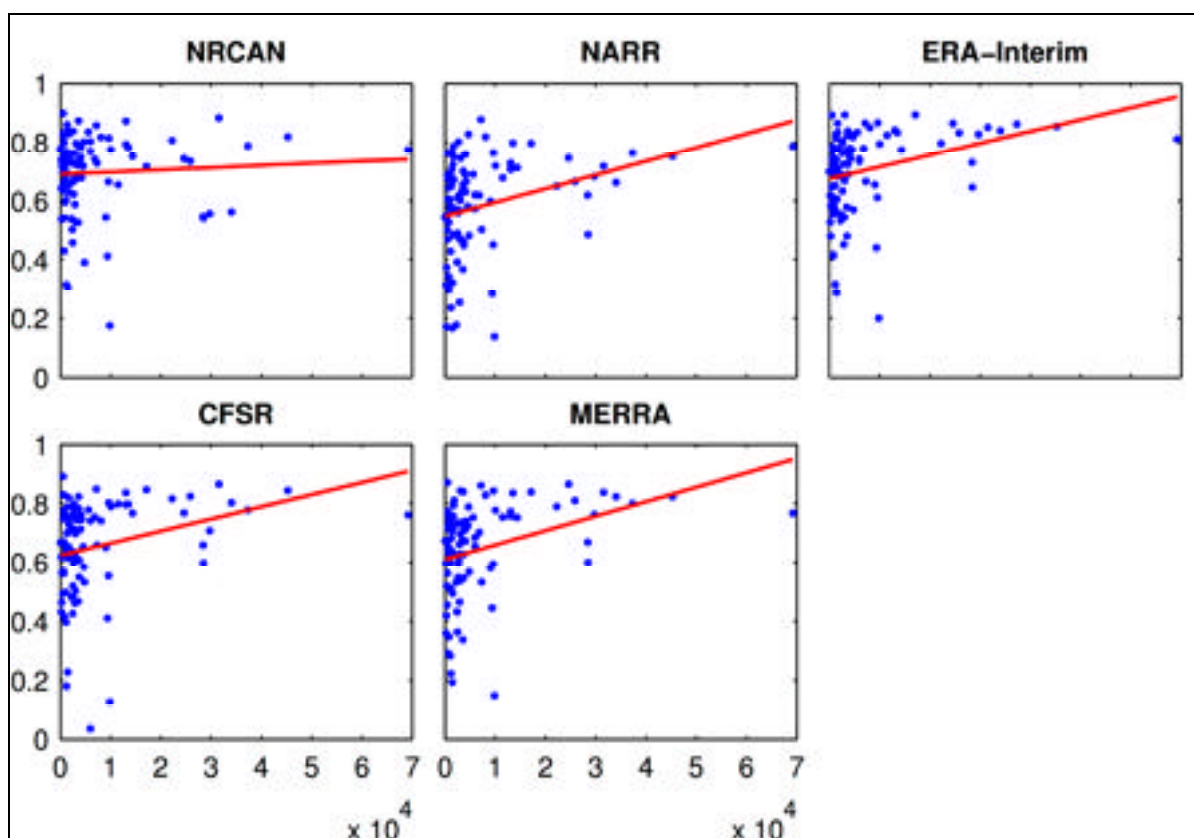


Figure 3.8 Distribution en nuage de points des coefficients de Nash-Sutcliffe (en ordonnée) des 99 bassins versants sélectionnés pour les jeux de données NRCAN, NARR, ERA-Interim, CFSR et MERRA. L'abscisse représente les superficies classées dans l'ordre croissant des bassins versants, allant de 10 à 69191 km² (de gauche à droite sur chaque sous-figure)

3.3 Discussions et conclusions

L'objectif de cette étude était d'évaluer la capacité des réanalyses NARR, ERA-Interim, CFSR et MERRA à reproduire les débits observés en rivière sur des bassins versants du Québec, notamment sur les bassins versants du nord, où les mesures de données météorologiques sont peu ou pas disponibles. Les performances des réanalyses à reproduire les débits observés en modélisation hydrologique sur 106 bassins versants de la base de données (cQ)² ont été comparées avec les performances du jeu de données d'observation sur grille NRCAN à l'aide du modèle global et conceptuel HSAMI, et les séries de débits

observés de ces bassins ont été utilisées comme référence. Le coefficient de Nash-Sutcliffe (NSE) a été utilisé pour évaluer la performance en calage et en validation des cinq jeux de données.

Bien que les différences de performances entre réanalyses ne soient systématiquement significatives en validation (voir figures en Annexe I), ERA-Interim se montre généralement supérieure aux trois autres réanalyses. MERRA et CFSR présentent des performances équivalentes, excepté sur la région Sud et en Gaspésie où CFSR domine légèrement MERRA. NARR présente des performances inférieures aux autres réanalyses, excepté dans la région Côte Nord où les performances de CFSR, NARR et MERRA sont sensiblement équivalentes. Malgré le fait que les résultats disponibles dans la région Arctique ne soient pas suffisants pour conclure sur la capacité des réanalyses à reproduire le débit des rivières dans cette région, les performances des réanalyses dans les autres régions sont prometteuses, notamment dans les régions nordiques, où le NSE peut s'élever au-dessus de 0,75, voire plus. Par ailleurs, les écarts entre les précipitations des différents jeux de données (Figure 2.6) sur certains bassins versants sont relativement atténués via le processus de calage du modèle HSAMI. En effet, en regardant par exemple les hydrogrammes moyens annuels des débits simulés par les réanalyses et NRCAN et des débits observés sur le bassin versant Baskatong (Figure 3.7), les cinq jeux de données révèlent des hydrogrammes relativement proches de celui des débits observés. De plus, le NSE moyen des cinq jeux de données sur ce bassin s'élève à 0,78 (Tableau 3.1). Ces résultats mettent en lumière la capacité des paramètres du modèle HSAMI à s'adapter, dans une certaine mesure, à différents types de précipitations, de manière à reproduire le plus fidèlement possible les débits observés.

Globalement, plus on descend dans le Sud, plus les performances des réanalyses sont faibles. D'autre part, l'étude de l'hydrogramme annuel moyen des débits à l'exutoire de douze bassins versants étudiés plus en détails révèle qu'il y a tout de même des écarts entre les débits simulés et observés, tantôt sur les étiages, tantôt sur les crues, et ces écarts s'accroissent du Nord au Sud. D'autre part, la résolution spatiale grossière des réanalyses comparée aux petites superficies des bassins versants dans le sud de la province pourrait aussi expliquer la

baisse de performance des réanalyses en modélisation hydrologique, bien que les résultats de cette étude n'aient pas démontré un lien particulièrement fort entre la performance des réanalyses et la superficie des bassins versants.

Néanmoins, les résultats de performance dépendent grandement de la qualité des séries de débits observés, comme cela a été démontré sur quatre bassins versants des régions Sud et Centre du Québec (Figure 3.5) où les performances des cinq jeux de données se révélaient être généralement médiocres. Malgré les intérêts que présente la base de données (cQ)², il n'en demeure pas moins que d'examiner la qualité des séries de débits observés demeure une étape préalable pertinente (voir section 3.1.1.1). Tel que mentionné à la section 3.2.1, il serait pertinent dans le cadre de futures études, de prendre en compte le critère d'homogénéité lors du processus de sélection des bassins versants.

Finalement, les réanalyses NARR, ERA-Interim, CFSR et MERRA peuvent se révéler utiles dans le cadre de simulations hydrologiques sur des bassins versants où les données météorologiques ne seraient pas disponibles. Toutefois, il serait intéressant d'approfondir l'évaluation des réanalyses en modélisation, notamment par l'application d'autres critères de performance, comme l'erreur quadratique moyenne ou l'erreur relative, ou par l'utilisation de modèles hydrologiques de nature différentes (distribués, à base physique). Enfin, l'analyse des simulations hydrologiques d'événements extrêmes par les réanalyses pourrait aussi se révéler très pertinente pour évaluer leur capacité à reproduire les conditions donnant lieu à de tels événements.

SYNTHÈSE ET CONCLUSION

L'évaluation du potentiel que représentent les ressources en eau douce du Québec nécessite de bien connaître le cycle hydrologique des différentes régions de cette province. Cependant, les données d'observation qui permettent de mieux comprendre et quantifier ce processus sont relativement rares dans certaines régions, notamment dans le nord de la province. Toutefois, il existe des jeux de données météorologiques alternatives aux données conventionnelles de stations, par exemple, les données interpolées sur grille, ou encore les réanalyses. Ces données alternatives ont la particularité de couvrir totalement un territoire et une période temporelle donnée, en fournissant des données sur une grille régulière.

L'objet de ce mémoire est d'évaluer, d'une part, le potentiel des réanalyses CFSR, ERA-Interim, MERRA et NARR à représenter le cycle hydrologique et, d'autre part, leur capacité à reproduire fidèlement les débits observés en modélisation hydrologique sur la province de Québec et ses bassins versants.

Dans un premier temps, le cycle hydrologique ainsi que ses composantes (précipitation, évaporation et ruissellement) ont été examinés durant la période 1979-2008 sur trois échelles spatiales différentes au Québec, et selon différentes échelles de temps: (1) moyenne à long terme des composantes hydrologiques sur la totalité du Québec, (2) cycle annuel moyen des composantes hydrologiques, et moyenne annuelle du bilan hydrologique sur les six régions climatiques du Québec selon la classification de Bukovsky (2011), et (3) moyennes annuelles de la précipitation et du ratio évaporation/précipitation sur 11 bassins versants du Québec, répartis sur six régions hydrologiques représentant les régimes hydriques majeurs de la province.

Les résultats de cette étude montrent que les quatre réanalyses ont leurs propres forces et limitations à représenter le cycle hydrologique au Québec. MERRA montre un bilan hydrologique équilibré par rapport aux trois autres réanalyses, mais son modèle de surface ne

répartit pas convenablement la précipitation en évaporation et en ruissellement. NARR, qui s'est révélée être assez fiable sur les États-Unis continentaux (voir notamment les détails sur les travaux de Gilles Essou, section 1.1.3), semble être la réanalyse la moins cohérente envers les observations sur grille et les autres réanalyses. En effet, de fortes sous-estimations ont été révélées dans la précipitation de NARR aux abords de la frontière canado-états-unienne, ainsi qu'entre les périodes 1979-2002 et 2003-2008 globalement sur le sud du Québec. MERRA et ERA-Interim présentent probablement les précipitations les plus fiables sur le Québec, alors que leurs évaporations et ruissellements demeurent discutables, à l'exception du ruissellement de ERA-Interim, qui présente des valeurs similaires aux observations. CFSR montre des précipitations surestimées par rapport aux trois autres réanalyses, mais présente vraisemblablement la meilleure répartition de la précipitation en évaporation et ruissellement, parmi les quatre réanalyses. Finalement, les réanalyses semblent globalement plus cohérentes entre elles dans le nord que dans le sud du Québec.

Dans un second temps, la performance en modélisation hydrologique de CFSR, ERA-Interim, MERRA et NARR a été évaluée sur 106 bassins versants du Québec, à l'aide du modèle conceptuel et global HSAMI. Le coefficient de Nash-Sutcliffe a été utilisé comme critère de performance en calage et en validation, et les résultats ont été analysés sur les mêmes régions hydrologiques définies lors de l'étude du cycle hydrologique des réanalyses.

Les résultats de cette étude montrent que les performances des réanalyses sont globalement satisfaisantes, avec un coefficient de Nash-Sutcliffe qui s'élève parfois au-dessus de 0,75. Les différences entre les quatre réanalyses ne sont pas statistiquement significatives, toutefois, ERA-Interim montre des performances globalement supérieures à celles des trois autres réanalyses. NARR présente les performances les plus basses parmi les quatre réanalyses, notamment dans les régions Sud et Centre, où l'étude de ses précipitations a révélé des brusques variations. Les réanalyses MERRA et CFSR font globalement jeu égal, excepté dans la région Sud et en Gaspésie. De plus, l'adaptation des paramètres du modèle HSAMI a permis au modèle de simuler, par le biais des réanalyses, des hydrogrammes moyens annuels

semblables à celui des débits observés, et des performances élevées sur des bassins versants où les réanalyses montraient des écarts importants dans leurs séries de précipitations.

Par ailleurs, les résultats de modélisation hydrologique peuvent être mis en relation avec les résultats des travaux de Gilles Essou en modélisation hydrologique avec ces quatre réanalyses sur des bassins versants des États-Unis (Essou et al., 2015). En effet, les performances des réanalyses diffèrent entre les bassins versants du Québec et des États-Unis. Parmi les quatre réanalyses, NARR, qui montre les performances les plus faibles au Québec, présente globalement les meilleurs résultats aux États-Unis. Rappelons que, outre la qualité variable des précipitations de NARR au Québec (évoqué précédemment), le réseau de données de précipitation assimilées dans NARR aux États-Unis est significativement plus dense qu'au Québec. Ce dernier aspect doit certainement contribuer aux bonnes performances de NARR sur les bassins versants des États-Unis. À l'inverse, les réanalyses ERA-Interim, MERRA et CFSR qui révèlent des performances globalement satisfaisantes au Québec, présentent des résultats moyens aux États-Unis (NSE entre 0,3 et 0,6), dont les bassins versants étudiés sont situés majoritairement dans l'Est du pays.

La précipitation sur grille NRCAN et la base de données de débits observés (cQ)² ont systématiquement été utilisées comme références lorsque cela était possible dans ces travaux. Pour cette raison, certains résultats doivent tout de même être mitigés, tant pour l'étude du cycle hydrologique que pour l'étude de la performance en modélisation. En effet, la précipitation de NRCAN et les débits observés de (cQ)² présentent des données de qualité variable, notamment à cause de la manière dont ces données ont été traitées: interpolation à partir d'un nombre trop faible de stations dans le nord pour NRCAN, et post-traitement des séries de débits pour (cQ)². De plus, dans le cas de l'interpolation des données NRCAN, le lissage au moyen des splines plaques minces peut affecter les valeurs obtenues sur tout le territoire du Québec, et donc altérer les valeurs des données aux stations. D'autre part, ces jeux de données révèlent parfois des erreurs de mesures qui n'ont pas été corrigées, ou ont été retirées dans le cas de (cQ)².

Néanmoins, les résultats n'en demeurent pas moins prometteurs, notamment sur les bassins des régions nordiques où les données aux stations se font rares: régions Nord et Côte Nord (60 % de la superficie totale des 106 bassins). En effet, les réanalyses y ont révélé des cycles hydrologiques somme toute cohérents, et de bonnes performances en modélisation hydrologiques. Des précautions devraient tout de même être prises avant l'utilisation des données de réanalyses sur le territoire du Québec, ou tout du moins, valider les données avant l'utilisation, notamment la précipitation de NARR. En effet, un modèle hydrologique conceptuel peut révéler une capacité d'adaptation malgré des éventuels biais par rapport aux observations présents dans les données de réanalyse.

Ainsi, il serait pertinent de valider les données de réanalyses par une comparaison directe avec les données des stations d'observation météorologiques. D'autre part, dans le cadre de futurs travaux en modélisation avec les réanalyses, l'utilisation d'autres types de modèles hydrologiques et d'autres critères de performance pourraient se révéler très intéressante et venir compléter les résultats présentés ici. De plus, l'étude des extrêmes, autant du cycle hydrologique que des débits simulés par les réanalyses, pourraient aussi révéler des informations pertinentes sur leur capacité à reproduire de tels événements.

D'autre part, après validation des données, les réanalyses pourraient être en mesure de fournir des informations supplémentaires aux modèles hydrologiques, comme l'évapotranspiration, la hauteur de neige ou l'humidité du sol. Ces variables pourraient aussi être utilisées comme données de référence pour évaluer les étapes intermédiaires de simulation, internes au modèle hydrologique (évapotranspiration, niveau de saturation du sol, fonte de neige), ou encore introduites directement dans les modèles hydrologiques pour chercher à réduire le nombre de paramètres de calage.

ANNEXE I

LISTE DES FIGURES RELATIVES AU TEST STATISTIQUE MULTIPLE DE COMPARAISON DE KRUSKAL-WALLIS EFFECTUÉ ENTRE LES DISTRIBUTIONS DES NSE DE CHAQUE JEU DE DONNÉES SUR LES SIX RÉGIONS HYDROLOGIQUES DU QUÉBEC

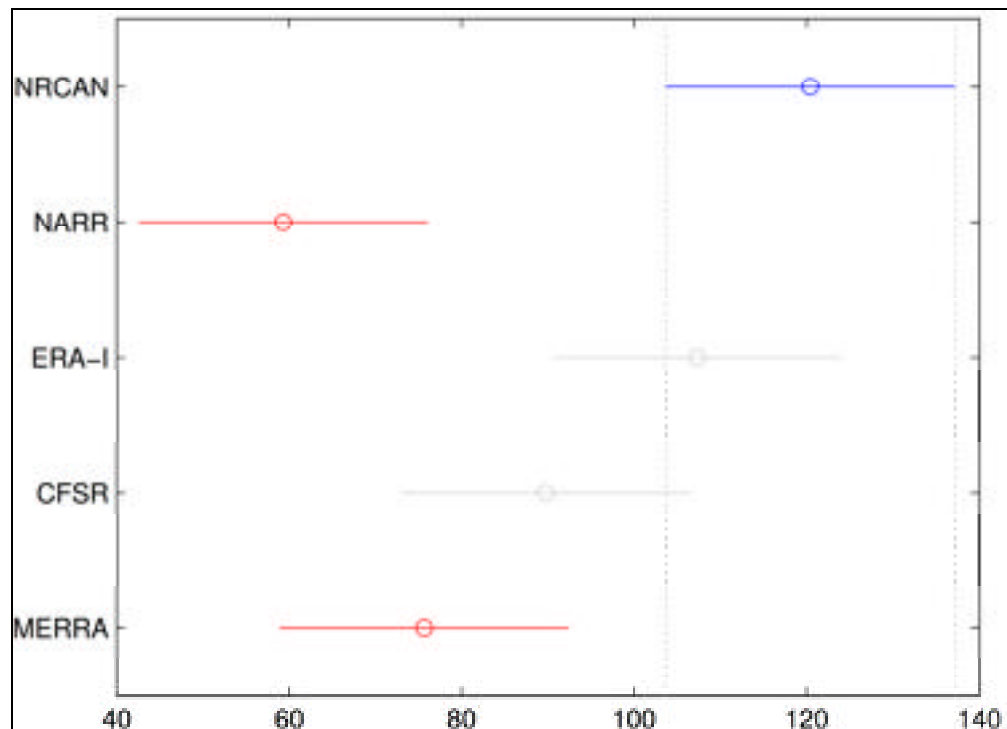


Figure-A I-1 Valeurs des rangs moyens (cercles) et de son intervalle de confiance (trait) des distributions de NSE en validation pour NRCAN, NARR, ERA-Interim, CFSR et MERRA sur la région **Sud**. En ordonnée: les noms des échantillons analysés (jeux de données), en abscisse: la valeur du rang des échantillons. Lorsque les couleurs des échantillons sont différentes, leurs distributions de NSE sont significativement différentes (les intervalles de confiance ne se chevauchent pas sur la figure)

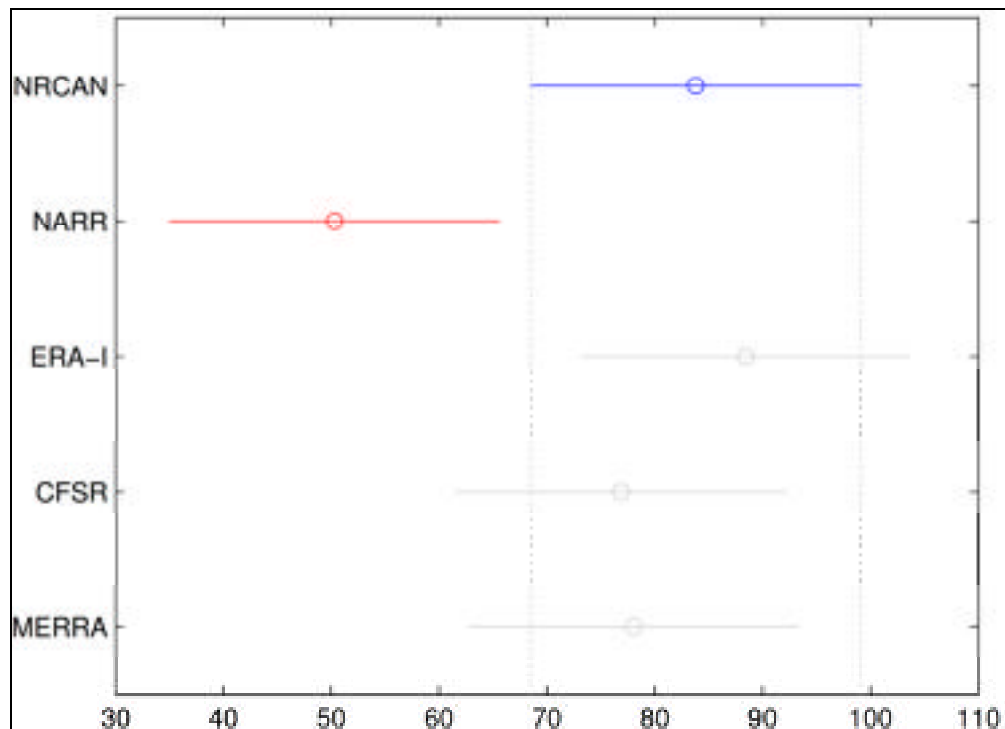


Figure-A I-2 Valeurs des rangs moyens (cercles) et de son intervalle de confiance (trait) des distributions de NSE en validation pour NRCAN, NARR, ERA-Interim, CFSR et MERRA sur la région **Centre**. En ordonnée: les noms des échantillons analysés (jeux de données), en abscisse: la valeur du rang des échantillons

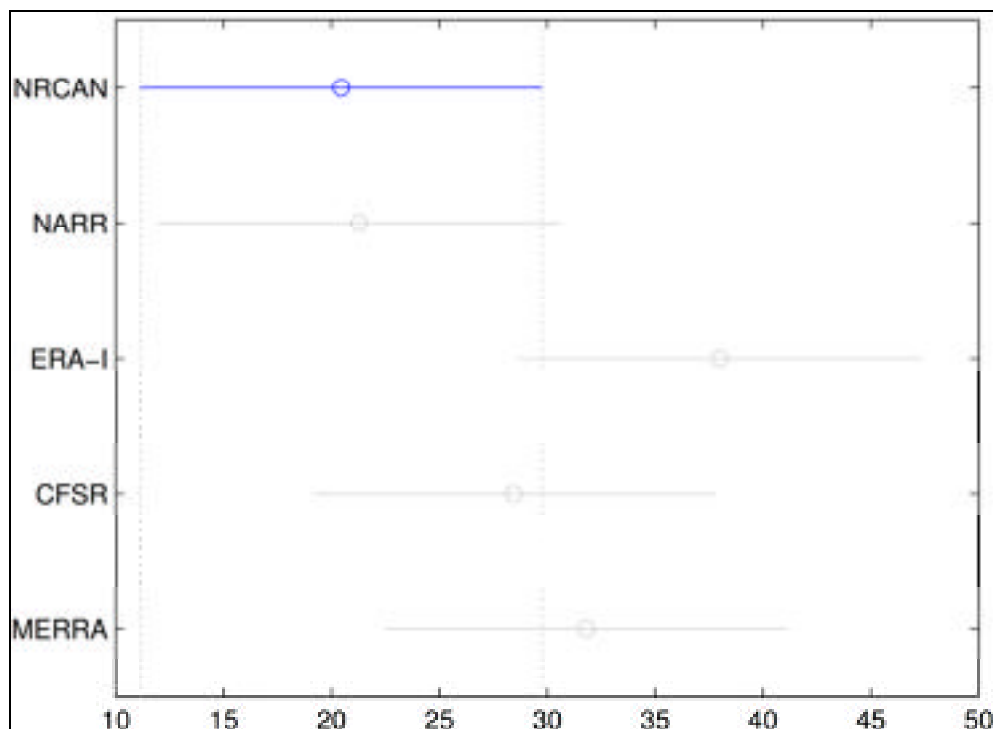


Figure-A I-3 Valeurs des rangs moyens (cercles) et de son intervalle de confiance (trait) des distributions de NSE en validation pour NRCAN, NARR, ERA-Interim, CFSR et MERRA sur la région **Nord**. En ordonnée: les noms des échantillons analysés (jeux de données), en abscisse: la valeur du rang des échantillons

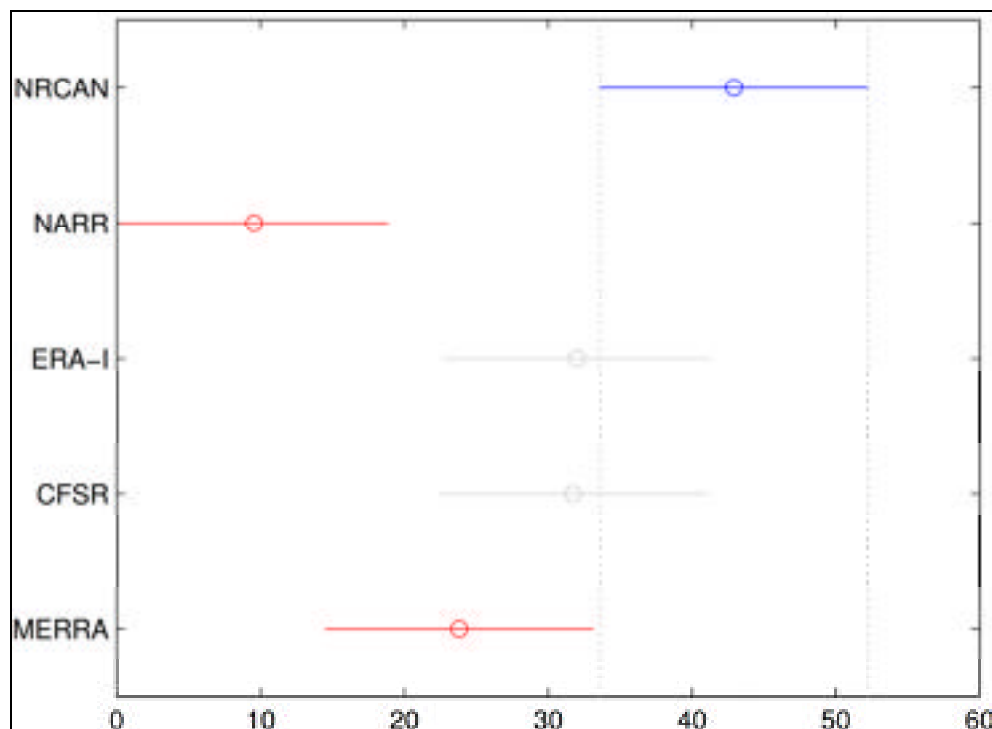


Figure-A I-4 Valeurs des rangs moyens (cercles) et de son intervalle de confiance (trait) des distributions de NSE en validation pour NRCAN, NARR, ERA-Interim, CFSR et MERRA sur la région **Gaspésie**. En ordonnée: les noms des échantillons analysés (jeux de données), en abscisse: la valeur du rang des échantillons

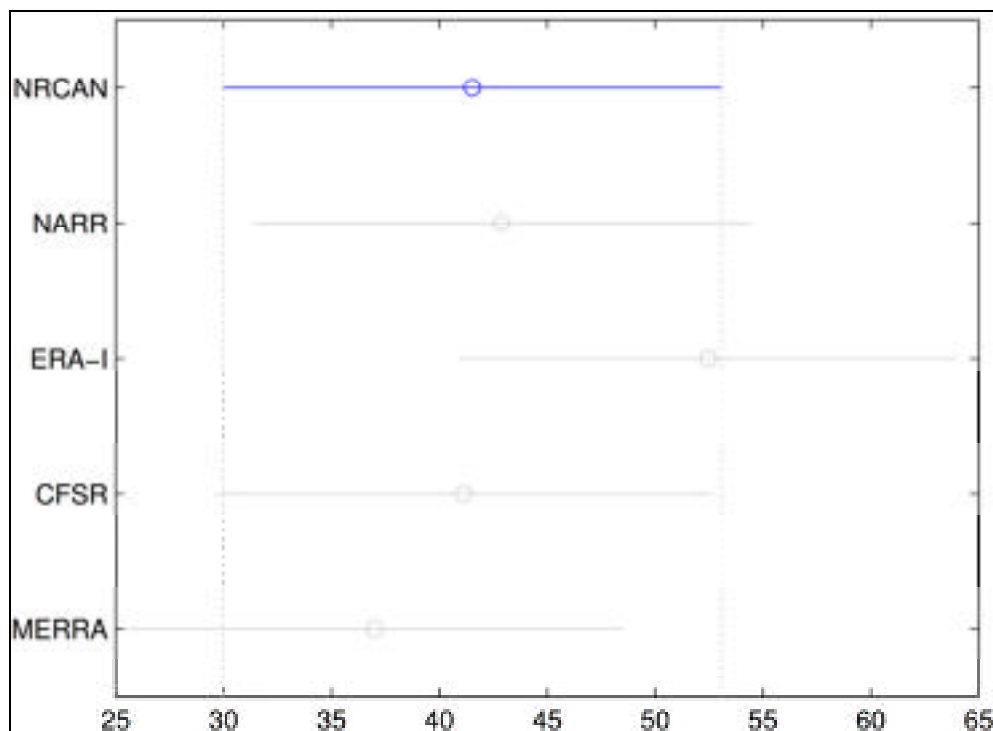


Figure-A I-5 Valeurs des rangs moyens (cercles) et de son intervalle de confiance (trait) des distributions de NSE en validation pour NRCAN, NARR, ERA-Interim, CFSR et MERRA sur la région **Côte Nord**. En ordonnée: les noms des échantillons analysés (jeux de données), en abscisse: la valeur du rang des échantillons

ANNEXE II

FIGURES DES MOYENNES SUR LA PÉRIODE 1979 - 2008 DES TEMPÉRATURES MINIMALES ET MAXIMALES DE NRCAN, CFSR, ERA-INTERIM ET MERRA SUR LES TERRITOIRES DE L'EST DU CANADA

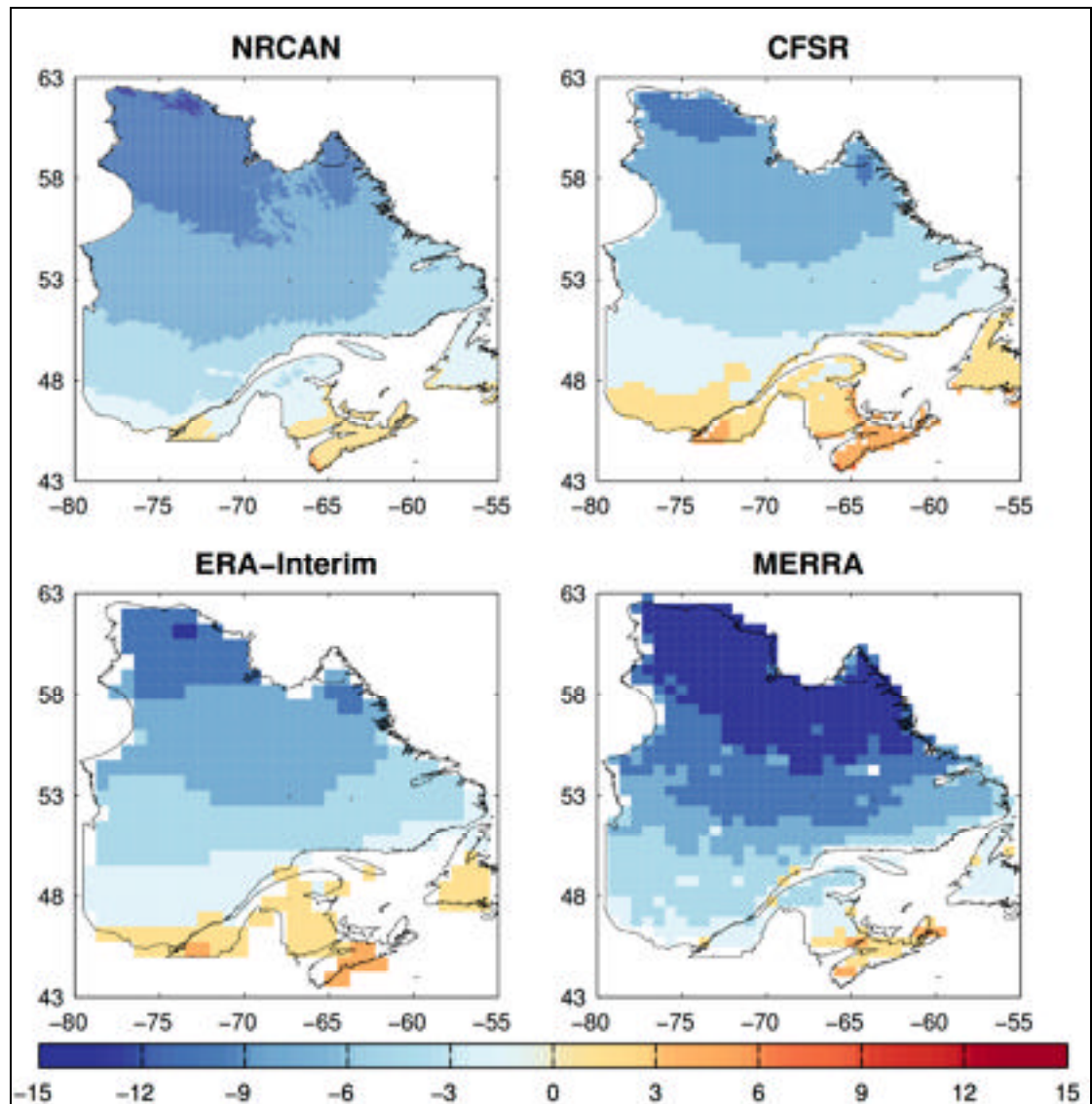


Figure-A II-1 Moyenne sur la période 1979 - 2008 des températures minimales de NRCAN, CFSR, ERA-Interim et MERRA sur les territoires de l'Est du Canada. Les températures sont en degré Celsius

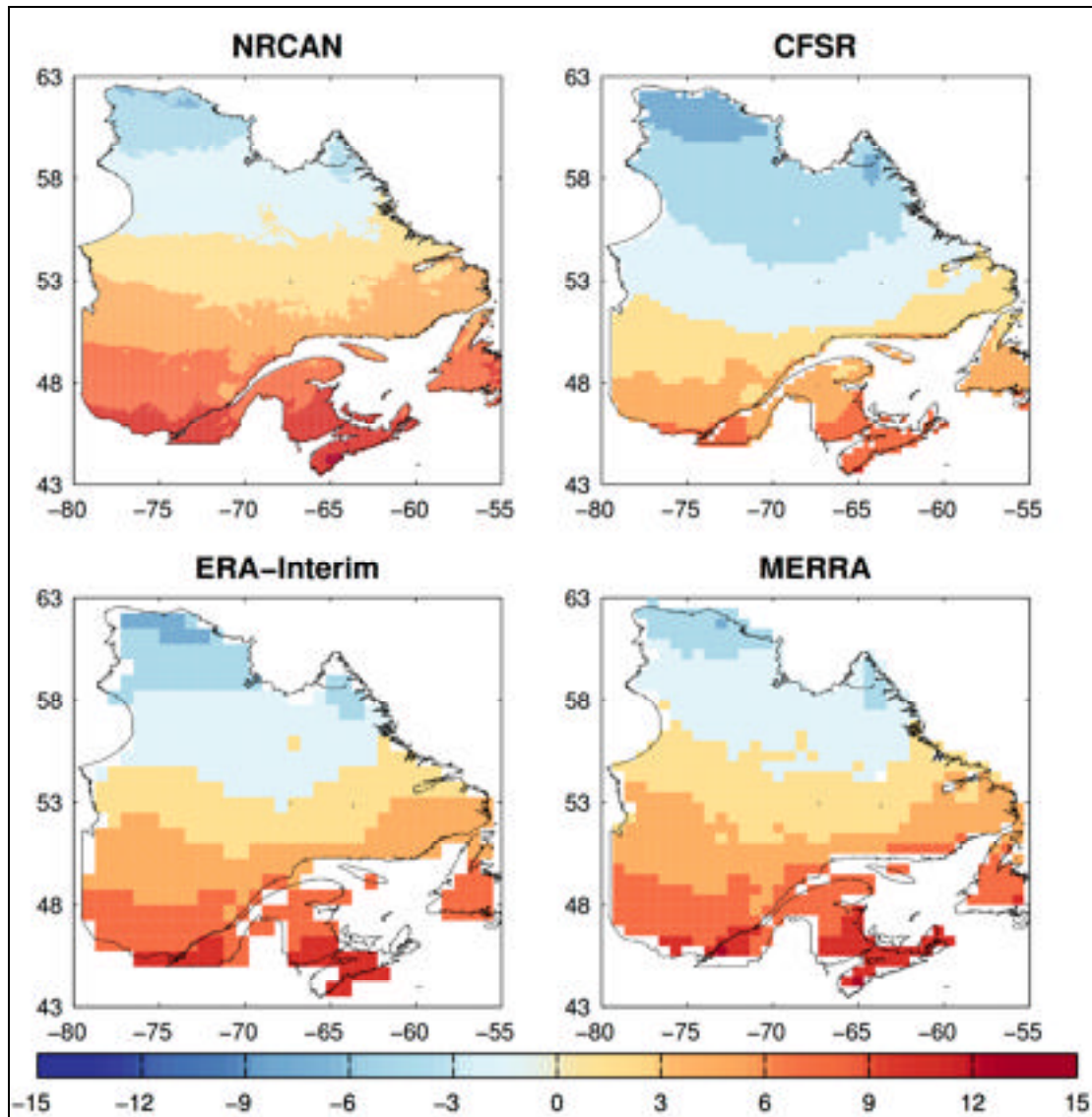


Figure-A II-2 Moyenne sur la période 1979 - 2008 des températures maximales de NRCAN, CFSR, ERA-Interim et MERRA sur les territoires de l'Est du Canada. Les températures sont en degré Celsius

LISTE DE RÉFÉRENCES BIBLIOGRAPHIQUES

- Arsenault, Richard, François Brissette, Jean-Stéphane Malo, Marie Minville et Robert Leconte. 2013a. « Structural and non-structural climate change adaptation strategies for the Péribonka water resource system ». *Water resources management*, vol. 27, n° 7, p. 2075-2087.
- Arsenault, Richard, et François P Brissette. 2014. « Continuous streamflow prediction in ungauged basins: The effects of equifinality and parameter set selection on uncertainty in regionalization approaches ». *Water Resources Research*, vol. 50, n° 7, p. 6135-6153.
- Arsenault, Richard, Annie Poulin, Pascal Côté et François Brissette. 2013b. « Comparison of stochastic optimization algorithms in hydrological model calibration ». *Journal of Hydrologic Engineering*, vol. 19, n° 7, p. 1374-1384.
- Association québécoise des techniques de l'eau, Office de la langue française. 1981. *Dictionnaire de l'eau*. Coll. « Cahiers de l'Office de la langue française ». Québec: Éditeur Officiel Du Québec, 544 p.
- Bengtsson, Lennart, et Jagadish Shukla. 1988. « Integration of space and in situ observations to study global climate change ». *Bulletin of the American Meteorological Society*, vol. 69, n° 10, p. 1130-1143.
- Betts, Alan K, Martin Köhler et Yuanchong Zhang. 2009. « Comparison of river basin hydrometeorology in ERA-Interim and ERA-40 reanalyses with observations ». *Journal of Geophysical Research: Atmospheres (1984–2012)*, vol. 114, n° D2.
- Beven, Keith. 2006. « A manifesto for the equifinality thesis ». In *The Model Parameter Estimation Experiment MOPEX*. 1-2 Vol. 320, p. 18-36. Coll. « Journal of Hydrology »: Elsevier. < <http://dx.doi.org/10.1016/j.jhydrol.2005.07.007> >.
- Bosilovich, Michael G, Junye Chen, Franklin R Robertson et Robert F Adler. 2008. « Evaluation of global precipitation in reanalyses ». *Journal of applied meteorology and climatology*, vol. 47, n° 9, p. 2279-2299.
- Bosilovich, Michael G, Franklin R Robertson et Junye Chen. 2011. « Global Energy and Water Budgets in MERRA ». *Journal of Climate*, vol. 24, n° 22, p. 5721–5739.
- Bukovsky, Melissa S. 2011. « Masks for the Bukovsky regionalization of North America ». < <http://www.narccap.ucar.edu/contrib/bukovsky/> >.

- Bukovsky, Melissa S., et David J. Karoly. 2007. « A brief evaluation of precipitation from the North American Regional Reanalysis ». *Journal of Hydrometeorology*, vol. 8, n° 4, p. 837-846.
- Canada, Fisheries and Environment. 1978. « Water balance - derived precipitation and evapotranspiration ». In *Hydrological Atlas of Canada*. < <http://geogratis.gc.ca/api/en/nrcan-rncan/ess-sst/910100c0-4f8c-5ae8-ae87-69bf230e43cf.html> >. Consulté le 2014-11-14.
- Chen, Jie, François P Brissette et Robert Leconte. 2011. « Uncertainty of downscaling method in quantifying the impact of climate change on hydrology ». *Journal of Hydrology*, vol. 401, n° 3, p. 190-202.
- Choi, Woonsup, Sung Joon Kim, Peter F. Rasmussen et Adam R. Moore. 2009. « Use of the North American regional reanalysis for hydrological modelling in Manitoba ». *Canadian Water Resources Journal*, vol. 34, n° 1, p. 17-36.
- Dee, DP, SM Uppala, AJ Simmons, P Berrisford, P Poli, S Kobayashi, U Andrae, MA Balmaseda, G Balsamo et P Bauer. 2011. « The ERA-Interim reanalysis: Configuration and performance of the data assimilation system ». *Quarterly Journal of the Royal Meteorological Society*, vol. 137, n° 656, p. 553-597.
- Dile, Yihun Taddele, et Raghavan Srinivasan. 2014. « Evaluation of CFSR climate data for hydrologic prediction in data-scarce watersheds: an application in the Blue Nile River Basin ». *JAWRA Journal of the American Water Resources Association*, vol. 50, n° 5, p. 1226-1241.
- Duan, Q, J Schaake, V Andreassian, S Franks, G Goteti, HV Gupta, YM Gusev, F Habets, A Hall et L Hay. 2006. « Model Parameter Estimation Experiment (MOPEX): An overview of science strategy and major results from the second and third workshops ». *Journal of Hydrology*, vol. 320, n° 1, p. 3-17.
- Essou, Gilles R. C., Florent Sabarly, Philippe Lucas-Picher, Francois Brissette et Annie Poulin. 2015. « Can precipitation and temperature from meteorological reanalyses be used for hydrological modeling? ». *Journal of Hydrometeorology*.
- Eum, Hyung-Il, Yonas Dibike, Terry Prowse et Barrie Bonsal. 2014. « Inter-comparison of high-resolution gridded climate data sets and their implication on hydrological model simulation over the Athabasca Watershed, Canada ». *Hydrological Processes*.
- Fortin, V. 2000. *Le Modèle Météo-Apport HSAMI: Historique, Théorie et Application*. Institut de recherche d'Hydro-Québec, 68 p.

- Fortin, Vincent, Guy Roy et Norman Donaldson. 2014. « Assimilation of radar QPE in the Canadian Precipitation Analysis (CaPA) ». In *ASCE International Symposium on Weather Radar and Hydrology, April 7th-10th, Washington*.
- Fuka, Daniel R, M Todd Walter, Charlotte MacAlister, Arthur T Degaetano, Tammo S Steenhuis et Zachary M Easton. 2014. « Using the Climate Forecast System Reanalysis as weather input data for watershed models ». *Hydrological Processes*, vol. 28, n° 22, p. 5613-5623.
- Gibson, JK. 1997. *ECMWF Re-Analysis Project report series: ERA description*, 1. European Centre for Medium-Range Weather Forecasts, 84 p.
- Gottschalck, Jon, Jesse Meng, Matt Rodell et Paul Houser. 2005. « Analysis of multiple precipitation products and preliminary assessment of their impact on global land data assimilation system land surface states ». *Journal of Hydrometeorology*, vol. 6, n° 5, p. 573-598.
- Hagemann, Stefan, et Lydia Dümenil Gates. 2001. « Validation of the hydrological cycle of ECMWF and NCEP reanalyses using the MPI hydrological discharge model ». *Journal of Geophysical Research: Atmospheres (1984–2012)*, vol. 106, n° D2, p. 1503-1510.
- Hansen, Nikolaus, et Andreas Ostermeier. 1996. « Adapting arbitrary normal mutation distributions in evolution strategies: The covariance matrix adaptation ». In *Evolutionary Computation, 1996., Proceedings of IEEE International Conference on*. p. 312-317. IEEE.
- Hansen, Nikolaus, et Andreas Ostermeier. 2001. « Completely derandomized self-adaptation in evolution strategies ». *Evolutionary computation*, vol. 9, n° 2, p. 159-195.
- Hutchinson, Michael F, Dan W McKenney, Kevin Lawrence, John H Pedlar, Ron F Hopkinson, Ewa Milewska et Pia Papadopol. 2009. « Development and testing of Canada-wide interpolated spatial models of daily minimum-maximum temperature and precipitation for 1961-2003 ». *Journal of Applied Meteorology and Climatology*, vol. 48, n° 4, p. 725-741.
- John M. Evans, Howard Perlman, USGS, Monica Michel. 2013. « Schéma du cycle de l'eau ». < <http://water.usgs.gov/edu/watercyclefrenchhi.html> >.
- Kalnay, E., M. Kanamitsu, R. Kistler, W. Collins, D. Deaven, L. Gandin, M. Iredell, S. Saha, C. White, J. Woollen, Y. Zhu, M. Chelliah, W. Ebisuzaki, W. Higgins, J. Janowiak, K. C. Mo, C. Ropelewski, J. Wang, A. Leetmaa, R. Reynolds, P. Jenne et D. Joseph. 1996. « The NCEP/NCAR 40-year reanalysis project ». *Bulletin of the American Meteorological Society*, vol. 77, n° 3, p. 437-71.

- Kanamitsu, Masao, Wesley Ebisuzaki, Jack Woollen, Shi-Keng Yang, JJ Hnilo, M Fiorino et GL Potter. 2002. « Ncep-doe amip-ii reanalysis (r-2) ». *Bulletin of the American Meteorological Society*, vol. 83, n° 11, p. 1631-1643.
- Kleidon, Axel, et Stan Schymanski. 2008. « Thermodynamics and optimality of the water budget on land: a review ». *Geophysical Research Letters*, vol. 35, n° 20, p. L20404.
- Kruskal, William H, et W Allen Wallis. 1952. « Use of ranks in one-criterion variance analysis ». *Journal of the American statistical Association*, vol. 47, n° 260, p. 583-621.
- Lauri, H, TA Räsänen et M Kummu. 2014. « Using Reanalysis and Remotely Sensed Temperature and Precipitation Data for Hydrological Modeling in Monsoon Climate: Mekong River Case Study ». *Journal of Hydrometeorology*, vol. 15, n° 4, p. 1532-1545.
- Lorenz, Christof, et Harald Kunstmann. 2012. « The hydrological cycle in three state-of-the-art reanalyses: intercomparison and performance analysis ». *Journal of Hydrometeorology*, vol. 13, n° 5, p. 1397-1420.
- Maurer, Edwin P, Greg M O'Donnell, Dennis P Lettenmaier et John O Roads. 2001. « Evaluation of the land surface water budget in NCEP/NCAR and NCEP/DOE reanalyses using an off-line hydrologic model ». *Journal of Geophysical Research: Atmospheres (1984–2012)*, vol. 106, n° D16, p. 17841-17862.
- Maurer, EP, AW Wood, JC Adam, DP Lettenmaier et B Nijssen. 2002. « A long-term hydrologically based dataset of land surface fluxes and states for the conterminous United States ». *Journal of Climate*, vol. 15, n° 22.
- Meng, Jesse, Rongqian Yang, Helin Wei, Michael Ek, George Gayno, Pingping Xie et Kenneth Mitchell. 2012. « The land surface analysis in the NCEP climate forecast system reanalysis ». *Journal of Hydrometeorology*, vol. 13, n° 5, p. 1621-1630.
- Mesinger, Fedor, Geoff DiMego, Eugenia Kalnay, Perry Shafran, Wesley Ebisuzaki, Dusan Jovic, Jack Woollen, Kenneth Mitchell, Eric Rogers, Michael Ek, Yun Fan, Robert Grumbine, Wayne Higgins, Hong Li, Ying Lin, Geoff Manikin, David Parrish et Wei Shi. 2004. « North American regional reanalysis ». In *Combined Preprints: 84th American Meteorological Society (AMS) Annual Meeting, January 11, 2004 - January 15, 2004*. (Seattle, WA., United states), p. 4189-4201. Coll. « Bulletin of the American Meteorological Society »: Amer. Meteor. Soc.
- Minville, Marie, François Brissette et Robert Leconte. 2009. « Impacts and uncertainty of climate change on water resource management of the Peribonka river system (Canada) ». *Journal of Water Resources Planning and Management*, vol. 136, n° 3, p. 376-385.

- Moradkhani, Hamid, et Soroosh Sorooshian. 2008. « General review of rainfall-runoff modeling: Model calibration, data assimilation, and uncertainty analysis ». In *Hydrological Modelling and the Water Cycle*. p. 1-24. Springer.
- Moriasi, DN, JG Arnold, MW Van Liew, RL Bingner, RD Harmel et TL Veith. 2007. « Model evaluation guidelines for systematic quantification of accuracy in watershed simulations ». *Trans. Asabe*, vol. 50, n° 3, p. 885-900.
- Onogi, Kazutoshi, Junichi Tsutsui, Hiroshi Koide, Masami Sakamoto, Shinya Kobayashi, Hiroaki Hatsushika, Takanori Matsumoto, Nobuo Yamazaki, Hirotaka KAMAHORI et Kiyotoshi Takahashi. 2007. « The JRA-25 reanalysis ». *Journal of the Meteorological Society of Japan*, vol. 85, n° 3, p. 369-432.
- Pechlivanidis, IG, BM Jackson, NR McIntyre et HS Wheeler. 2011. « Catchment scale hydrological modelling: a review of model types, calibration approaches and uncertainty analysis methods in the context of recent developments in technology and applications ». *Global NEST journal*, vol. 13, n° 3, p. 193-214.
- Peixoto, José P, et Abraham H Oort. 1992. « Physics of climate ».
- Poulin, Annie, François Brissette, Robert Leconte, Richard Arsenault et Jean-Stéphane Malo. 2011. « Uncertainty of hydrological modelling in climate change impact studies in a Canadian, snow-dominated river basin ». *Journal of Hydrology*, vol. 409, n° 3, p. 626-636.
- Pushpalatha, Raji, Charles Perrin, Nicolas Le Moine et Vazken Andréassian. 2012. « A review of efficiency criteria suitable for evaluating low-flow simulations ». *Journal of Hydrology*, vol. 420, p. 171-182.
- Rhynsburger, Dierk. 1973. « Analytic Delineation of Thiessen Polygons* ». *Geographical Analysis*, vol. 5, n° 2, p. 133-144.
- Riboust, Philippe. 2014. *Modélisation et analyse de l'impact du changement climatique sur les inondations de la rivière du Richelieu mises en contexte avec la crue extrême du printemps 2011* (2014). Montréal: École de technologie supérieure, 1 ressource en ligne (xviii, 120 pages) p.
- Ricketts, Taylor H, Eric Dinerstein, David M Olson et Colby J Loucks. 1999. *Terrestrial ecoregions of North America: a conservation assessment*, 1. Washington, DC: Island Press, 485 p.
- Rienecker, Michele M, Max J Suarez, Ronald Gelaro, Ricardo Todling, Julio Bacmeister, Emily Liu, Michael G Bosilovich, Siegfried D Schubert, Lawrence Takacs et Gi-

- Kong Kim. 2011. « MERRA: NASA's Modern-Era Retrospective Analysis for Research and Applications ». *Journal of Climate*, vol. 24, n° 14, p. 3624–3648.
- Roads, JO, S-C Chen, M Kanamitsu et H Juang. 1998. « Vertical structure of humidity and temperature budget residuals over the Mississippi River basin ». *Journal of Geophysical Research: Atmospheres (1984–2012)*, vol. 103, n° D4, p. 3741-3759.
- Saha, Suranjana, Shrinivas Moorthi, Hua-Lu Pan, Xingren Wu, Jiande Wang, Sudhir Nadiga, Patrick Tripp, Robert Kistler, John Woollen et David Behringer. 2010. « The NCEP climate forecast system reanalysis ». *Bulletin of the American Meteorological Society*, vol. 91, n° 8, p. 1015-1057.
- Sheffield, Justin, Ben Livneh et Eric F Wood. 2012. « Representation of terrestrial hydrology and large-scale drought of the continental United States from the North American Regional Reanalysis ». *Journal of Hydrometeorology*, vol. 13, n° 3, p. 856-876.
- Trenberth, Kevin E, John T Fasullo et Jessica Mackaro. 2011. « Atmospheric moisture transports from ocean to land and global energy flows in reanalyses ». *Journal of Climate*, vol. 24, n° 18, p. 4907-4924.
- Uppala, S. M., P. W. Kallberg, Adrian J. Simmons, U. Andrae, V. da Costa Bechtold, M. Fiorino, J. K. Gibson, J. Haseler, A. Hernandez, G. A. Kelly, X. Li, K. Onogi, S. Saarinen, N. Sokka, R. P. Allan, E. Andersson, K. Arpe, M. A. Balmaseda, A. C. M. Beljaars, L. van de Berg, J. Bidlot, N. Bormann, S. Caires, F. Chevallier, A. Dethof, M. Dragosavac, M. Fisher, M. Fuentes, S. Hagemann, E. Holm, B. J. Hoskins, L. Isaksen, P. A. E. M. Janssen, R. Jenne, A. P. McNally, J. F. Mahfouf, J. J. Morcrette, N. A. Rayner, R. W. Saunders, P. Simon, A. Sterl, K. E. Trenberth, A. Untch, D. Vasiljevic, P. Viterbo et J. Woollen. 2005. « The ERA-40 re-analysis ». *Quarterly Journal of the Royal Meteorological Society*, vol. 131, n° 612, p. 2961-3012.
- Wang, Shusen, Daniel W McKenney, Jiali Shang et Junhua Li. 2014. « A national-scale assessment of long-term water budget closures for Canada's watersheds ». *Journal of Geophysical Research: Atmospheres*, vol. 119, n° 14, p. 8712-8725.
- Wang, Wanqiu, Pingping Xie, Soo-Hyun Yoo, Yan Xue, Arun Kumar et Xingren Wu. 2011. « An assessment of the surface climate in the NCEP climate forecast system reanalysis ». *Climate dynamics*, vol. 37, n° 7-8, p. 1601-1620.
- Zhang, Li, Arun Kumar et Wanqiu Wang. 2012. « Influence of changes in observations on precipitation: A case study for the Climate Forecast System Reanalysis (CFSR) ». *Journal of Geophysical Research: Atmospheres (1984–2012)*, vol. 117, n° D8.